GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 18:17:02 ; Search time 1853.5 Seconds
(without alignments)
1240.565 Million cell updates/sec

US-10-090-326-25
77
1 acttacncagcccagcanca.....ggagnacaacagccancagt 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

27513289 seqs, 14931090276 residues

hits satisfying chosen parameters: 55026578

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

is the number of results predicted by chance to have a
ıter than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

	% Query Match	Length	DB	ID	Description
	49.1	939	12	BG764646	BG764646 602736726
	39.2	522	28	AQ585695	AQ585695 RPCI-11-4
	39.0	294	13	BQ678960	BQ678960 AGENCOURT
	39.0	390	14	N28965	N28965 yx39d02.r1
	39.0	674	13	BX102971	BX1029
	39.0	764	12	BI859609	BI8596
	39.0	908	13	BQ424713	BQ4247
	39.0	926	13	BQ424642	BQ4246
	39.0	943	13	BQ677071	BQ6770
	39.0	965	12	BG576440	BG5764
	38.7	1063	12	BM463960	BM4639
c	37.9	1155	13	BX340723	BX3407
c	37.4	461	28	AZ095479	AZ0954
	37.4	651	28	AZ661932	AZ6619
c	37.1	935	10	BF237982	BF2379
c	36.6	400	13	BX749935	BX7499
	36.4	617	28	AZ658637	AZ6586
c	36.4	625	29	CE510405	CE5104
c	36.4	939	13	BU465668	BU4656
c	35.8	544	28	AZ135044	AZ13504
c	35.8	781	28	AQ860588	AQ8605
c	35.6	221	10	BB588106	BB5881
c	35.6	301	9	AU183092	AU183092
	35.6	304	29	CE670554	CE6705
c	35.6	330	9	AU182930	AU182930
c	35.6	447	9	AU183801	AU183801
c	35.6	451	9	AU223051	AU223051
c	35.6	453	9	AU222931	AU222931
c	35.6	454	9	AU183799	AU183799
c	35.6	470	9	AU172933	AU172933
c	35.6	482	12	BM487714	BM48771
	35.6	687	14	CB685142	CB6851
	35.6	782	29	CE474611	CE4746
	35.6	984	13	BU190117	BU1901
	35.6	1068	12	BM557397	BM5573
c	35.3	462	12	BM595860	BM5958
c	35.3	465	12	BM644507	BM6445
c	35.3	481	12	BM648778	BM6487
c	35.3	482	11	CNS08F9W	BX0092
c	35.3	487	12	BM643718	BM6437
c	35.3	503	12	BM655292	BM6552
c	35.3	507	12	BM622186	BM6221
c	35.3	510	12	BM620575	BM6205
c	35.3	512	12	BM587290	BM5872
c	35.3	541	12	BM644366	BM6443

ALIGNMENTS

RESULT 1
BG764646
LOCUS BG764646 939 bp mRNA linear EST
DEFINITION 602736726F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:48
mRNA sequence.
ACCESSION BG764646
VERSION BG764646.1 GI:14075299
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 939)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collectio
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LL
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1721 row: d column: 07
High quality sequence stop: 885.

```
       Location/Qualifiers
       1. .939
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:4861782"
       /tissue_type="melanotic melanoma, high MDR (cell line)"
       /lab_host="DH10B (phage-resistant)"
       /clone_lib="NIH_MGC_49"
       /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
       EcoRI; cDNA made by oligo-dT priming. Directionally cloned
       into EcoRI/XhoI sites using the following 5' adaptor:
       GGCACGAG(G). Size-selected >500bp for average insert size
       1.8kb. Library constructed by Ling Hong in the laboratory
       of Gerald M. Rubin (University of California, Berkeley)
       using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
       II RT (Life Technologies). Note: this is a NIH_MGC
       Library. |"


            49.1%;  Score 37.8;  DB 12;  Length 939;
imilarity   66.2%;  Pred. No. 0.077;
;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
'|||||  ||||||||||  ||||||       |   | |     |   | |  ||||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATGTCTGTAGCCGATTGGAGGA 326

ENACAACAGCCANCAGT 77
 ||||||||||  ||||
GTACAACAGCCATCAGT 343



385695                 522 bp    DNA     linear   GSS 07-JUN-1999
CI-11-435G16.TJ RPCI-11 Homo sapiens genomic clone
CI-11-435G16, genomic survey sequence.
385695
385695.1  GI:5012741
).
lo sapiens (human)
lo sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 522)
lo,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
ater,J.C.
e of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Building
ublished (1997)
ier_GSSs: RPCI-11-435G16.TV
ntact: Shaying Zhao, William Nierman, Mark Adams
partment of Eukaryotic Genomics
e Institute for Genomic Research
12 Medical Center Dr., Rockville, MD 20850
l: 301 838 0200
x: 301 838 0208
ail: hbe@tigr.org
nes are derived from the human BAC library RPCI-11. For BAC
rary availability, please contact Pieter de Jong
ieter@dejong.med.buffalo.edu). Clones may be purchased from
CPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
esearch Genet cs (info@resgen.com). BAC end search page:
tp://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 primer: SP6
ass: BAC ends.
       Location/Qualifiers
       1. .522
       /organism="Homo sapiens"
       /mol_type="genomic DNA"
       /db_xref="GDB:7666815"
       /db_xref="taxon:9606"
       /clone="RPCI-11-435G16"
       /sex="Male"
       /cell_type="Lymphocytes"
       /clone_lib="RPCI-11"
       /note="Vector: pBACe3.6; Site_1: EcoRI; Site_
       RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             39.2%;  Score 30.2;  DB 28;  Length 522;
  Best Local Similarity   62.0%;  Pred. No. 20;
  Matches   44;  Conservative    0;  Mismatches   27;  Indels    0

Qy          5 ACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGG
              || |||||  ||   ||| |  | |  || |   | | |  | | || ||
Db        485 ACACAGCCAGGCTAGATTATCTCGTTGTCATAATCTTTGCAAAGGCAGTTTC

Qy         65 AACAGCCANCA 75
              |||||||| ||
Db        425 AACAGCCACCA 415


RESULT 3
BQ678960
LOCUS       BQ678960                 294 bp    mRNA    linear   EST
DEFINITION  AGENCOURT_8281522 NIH_MGC_112 Homo sapiens cDNA clone
            5', mRNA sequence.
ACCESSION   BQ678960
VERSION     BQ678960.1  GI:21791639
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 294)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collectic
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LI
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution informatio
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2423  row: i  column: 24
            High quality sequence stop: 293.
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6261983"
                     /tissue_type="melanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_112"
                     /note="Organ: skin; Vector: pOTB7; Site_1: Xho
                     EcoRI; cDNA made by oligo-dT priming. Directio
                     into EcoRI/XhoI sites using the following 5' a
                     GGCACGAG(G). Library constructed  by Ling Hong
                     laboratory of Gerald M. Rubin (University of
                     Berkeley) using ZAP-cDNA synthesis kit (Strata
                     Superscript II RT (Life Technologies). Note: t
                     NIH_MGC Library."
ORIGIN

  Query Match             39.0%;  Score 30;  DB 13;  Length 294;
  Best Local Similarity   67.9%;  Pred. No. 18;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
```

```
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 268

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 286
```

```
8965                    390 bp    mRNA    linear   EST 04-JAN-1996
39d02.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
AGE:264099 5' similar to gb:M27160_rna1 TYROSINASE PRECURSOR
UMAN);, mRNA sequence.
8965
8965.1  GI:1147201
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 390)
llier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
lman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
rsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
evaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
lson,R.
e WashU-Merck EST Project
published (1995)
ntact: Wilson RK
shington University School of Medicine
44 Forest Park Parkway, Box 8501, St. Louis, MO 63108
l: 314 286 1800
x: 314 286 1810
ail: est@watson.wustl.edu
gh quality sequence stops: 296
urce: IMAGE Consortium, LLNL
is clone is available royalty-free through LLNL ; contact the
AGE Consortium (info@image.llnl.gov) for further information.
q primer: T7
gh quality sequence stop: 296.
       Location/Qualifiers
       1. .390
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="GDB:3873741"
       /db_xref="taxon:9606"
       /clone="IMAGE:264099"
       /sex="Male"
       /tissue_type="melanocyte"
       /lab_host="DH10B (ampicillin resistant)"
       /clone_lib="Soares melanocyte 2NbHM"
       /note="Vector: pT7T3D (Pharmacia) with a modified
       polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
       was primed with a Not I - oligo(dT) primer [5'
       TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTT 3'],
       double-stranded cDNA was size selected, ligated to Eco RI
       adapters (Pharmacia), digested with Not I and cloned into
       the Not I and Eco RI sites of a modified pT7T3 vector
       (Pharmacia). Library constructed by Bento Soares and
       M.Fatima Bonaldo. RNA from normal foreskin melanocytes
       (FS374) was kindly provided by Dr. Anthony P. Albino."


           39.0%;  Score 30;  DB 14;  Length 390;
imilarity   67.9%;  Pred. No. 21;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 158

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
```

```
Db        159 AGTACAACAGCCATCAGT 176


RESULT 5
BX102971
LOCUS       BX102971                674 bp    mRNA    linear   E
DEFINITION  BX102971 Soares melanocyte 2NbHM Homo sapiens cDNA cl
            IMAGp998B04578 ; IMAGE:264099, mRNA sequence.
ACCESSION   BX102971
VERSION     BX102971.1  GI:27832125
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 674)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung G
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998B04578.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rol
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung G
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information.
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .674
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998B04578 ; IMAGE:264099"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /note="Vector: pT7T3D (Pharmacia) with a modi
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1s
                     was primed with a Not I - oligo(dT) primer [5
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTT
                     double-stranded cDNA was size selected, ligat
                     adapters (Pharmacia), digested with Not I and
                     the Not I and Eco RI sites of a modified pT7T
                     (Pharmacia). Library constructed by Bento Soa
                     M.Fatima Bonaldo. RNA from normal foreskin me
                     (FS374) was kindly provided by Dr. Anthony P.
ORIGIN

  Query Match           39.0%;  Score 30;  DB 13;  Length 674;
  Best Local Similarity   67.9%;  Pred. No. 27;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1
```

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        127 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        187 AGTACAACAGCCATCAGT 204


RESULT 6
BI859609
```

59609 764 bp mRNA linear EST 10-OCT-2001
388079F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5396923 5',
A sequence.
59609
59609.1 GI:16000356
.
io sapiens (human)
io sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 764)
H-MGC http://mgc.nci.nih.gov/.
tional Institutes of Health, Mammalian Gene Collection (MGC)
ublished (1999)
ntact: Robert Strausberg, Ph.D.
ail: cgapbs-r@mail.nih.gov
ssue Procurement: DCTD/DTP
DNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
NA Sequencing by: Incyte Genomics, Inc.
lone distribution: MGC clone distribution information can be
nd through the I.M.A.G.E. Consortium/LLNL at:
p://image.llnl.gov
te: LLAM12011 row: m column: 20
gh quality sequence stop: 763.
Location/Qualifiers
1. .764
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5396923"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

39.0%; Score 30; DB 12; Length 764;
imilarity 67.9%; Pred. No. 28;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| |||||| | | || | | ||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 543

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 561

424713 908 bp mRNA linear EST 23-MAY-2002
ENCOURT_7926435 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6160418
, mRNA sequence.
424713
424713.1 GI:21120028
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 908)
H-MGC http://mgc.nci.nih.gov/.
tional Institutes of Health, Mammalian Gene Collection (MGC)
published (1999)
ntact: Robert Strausberg, Ph.D.
ail: cgapbs-r@mail.nih.gov
ssue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LL
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution informatior
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13511 row: b column: 03
High quality sequence stop: 628.
FEATURES Location/Qualifiers
source 1. .908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6160418"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_
Site_2: SalI; Cloned unidirectionally. Primer
Average insert size 2 kb. Library constructe
Technologies."
ORIGIN

Query Match 39.0%; Score 30; DB 13; Length 908;
Best Local Similarity 67.9%; Pred. No. 30;
Matches 53; Conservative 0; Mismatches 24; Indels 1,

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | || | | ||||||
Db 445 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 505 AGTACAACAGCCATCAGT 522

RESULT 8
BQ424642
LOCUS BQ424642 926 bp mRNA linear ES
DEFINITION AGENCOURT_7926673 NIH_MGC_72 Homo sapiens cDNA clone I
5', mRNA sequence.
ACCESSION BQ424642
VERSION BQ424642.1 GI:21119957
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE 1 (bases 1 to 926)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collecti
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (L
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution informatio
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13511 row: a column: 18
High quality sequence stop: 628.
FEATURES Location/Qualifiers
source 1. .926
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6160409"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

39.0%; Score 30; DB 13; Length 926;
milarity 67.9%; Pred. No. 31;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| |||||| | | | || | | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 501

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 519

77071 943 bp mRNA linear EST 15-JUL-2002
NCOURT_8209637 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6272720
mRNA sequence.
77071
77071.1 GI:21789750
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 943)
-MGC http://mgc.nci.nih.gov/.
ional Institutes of Health, Mammalian Gene Collection (MGC)
iblished (1999)
tact: Robert Strausberg, Ph.D.
il: cgapbs-r@mail.nih.gov
sue Procurement: DCTD/DTP
NA Library Preparation: Rubin Laboratory
NA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
A Sequencing by: Agencourt Bioscience Corporation
one distribution: MGC clone distribution information can be
id through the I.M.A.G.E. Consortium/LLNL at:
p://image.llnl.gov
e: LLCM2451 row: i column: 09
a quality sequence stop: 636.
Location/Qualifiers
1. .943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6272720"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

39.0%; Score 30; DB 13; Length 943;
ilarity 67.9%; Pred. No. 31;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||| |||||||||| |||||| | | | || | | | |||||||
TTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 267

NACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 268 AGTACAACAGCCATCAGT 285

RESULT 10
BG576440
LOCUS BG576440 965 bp mRNA linear EST
DEFINITION 602595683F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:47
mRNA sequence.
ACCESSION BG576440
VERSION BG576440.1 GI:13584093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collectio
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LL
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10580 row: f column: 22
High quality sequence stop: 737.
FEATURES Location/Qualifiers
source 1. .965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4708629"
/tissue_type="mammary adenocarcinoma, cell lin
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/note="Organ: breast; Vector: pCMV-SPORT6; Sit
Site_2: SalI; Cloned unidirectionally; oligo-d
Average insert size 1.383 kb. Library enriched
full-length clones and constructed by Life Tec
Note: this is a NIH_MGC Library."
ORIGIN

Query Match 39.0%; Score 30; DB 12; Length 965;
Best Local Similarity 67.9%; Pred. No. 31;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
|||||| |||||||||| |||||| | | | || | | |
Db 92 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 152 AGTACAACAGCCATCAGT 169

RESULT 11
BM463960
LOCUS BM463960 1063 bp mRNA linear EST
DEFINITION AGENCOURT_6445513 NIH_MGC_72 Homo sapiens cDNA clone IM
5', mRNA sequence.
ACCESSION BM463960
VERSION BM463960.1 GI:18513002
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
REFERENCE 1 (bases 1 to 1063)

```
-MGC http://mgc.nci.nih.gov/.
ional Institutes of Health, Mammalian Gene Collection (MGC)
ublished (1999)
tact: Robert Strausberg, Ph.D.
il: cgapbs-r@mail.nih.gov
sue Procurement: ATCC/DCTD/DTP
NA Library Preparation: Life Technologies, Inc.
NA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
A Sequencing by: Agencourt Bioscience Corporation
one distribution: MGC clone distribution information can be
nd through the I.M.A.G.E. Consortium/LLNL at:
p://image.llnl.gov
te: LLAM12235  row: f  column: 10
h quality sequence stop: 749.
      Location/Qualifiers
      1. .1063
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5539977"
      /tissue_type="melanotic melanoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_72"
      /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
      Average insert size 2 kb.   Library constructed by Life
      Technologies."


          38.7%;  Score 29.8;  DB 12;  Length 1063;
milarity   59.7%;  Pred. No. 38;
  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
||||| ||| || ||| ||||       |   | |     |  | |   |||||||
CTTACTCAGGCCCGCATCATTTCTCTCCTCTTGGCAGATGTCTGTAGCCCGATGGAGGA 953

NACAACAGCCANCAGT 77
  ||||||||||| ||||
TACAACAGCCATCAGT 970



40723                1155 bp    mRNA    linear   EST 01-MAY-2003
40723 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
o sapiens cDNA clone CS0DJ005YO08 5-PRIME, mRNA sequence.
40723
40723.1  GI:30309280
'.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 1155)
W.B., Gruber,C., Jessee,J. and Polayes,D.
l-length cDNA libraries and normalization
ublished (2001)
tact: Genoscope
oscope - Centre National de Sequencage
191 91006 EVRY cedex - France
il: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
rary was constructed by Life Technologies, a division of
itrogen. This sequence belongs to sequence cluster 10466.f For
e information about this cluster, see
p://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ005BH04QP1
uster=10466.f. Contact : Feng Liang Email : fliang@lifetech.com
 : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
aday Avenue Genoscope sequence ID : CS0DJ005BH04QP1.
      Location/Qualifiers
      1. .1155
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DJ005YO08"
      /cell_type="T CELLS (JURKAT CELL LINE) COT 10-
      /cell_line="JURKAT"
      /clone_lib="Homo sapiens T CELLS (JURKAT CELL
      10-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-
      primer. Five prime end enriched, double-strand
      digested with Not I and cloned into the Not I
      sites of the pCMVSPORT 6 vector. Library was n
ORIGIN

  Query Match             37.9%;  Score 29.2;  DB 13;  Length 1155;
  Best Local Similarity   52.8%;  Pred. No. 63;
  Matches   38;  Conservative    8;  Mismatches   26;  Indels    0;

Qy          5 ACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGA
              ||  :  |||||| | :||| || :|||  || : |   :|| |||| | ||
Db       1052 ACCAVCACCAGCACCTKTCTGGCCYCATTTTCASAAATYYCATGTCAGTGGGK

Qy         65 AACAGCCANCAG 76
              | :|| :| ||:
Db        992 AGSAGRSAGCAR 981


RESULT 13
AZ095479/c
LOCUS       AZ095479                 461 bp    DNA     linear   GSS
DEFINITION  RPCI-23-474A7.TJ RPCI-23 Mus musculus genomic clone RPC
            genomic survey sequence.
ACCESSION   AZ095479
VERSION     AZ095479.1  GI:7737522
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K.,
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-474A7.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23.
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchase
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderin
            or from Resea ch Genetics (info@resgen.com). BAC end pa
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.ht
            Plate: 474  row: A  column: 7
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-474A7"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6;
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse ki
                     brain genomic DNA was isolated and  partially
```

with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). "

37.4%; Score 28.8; DB 28; Length 461;
milarity 62.7%; Pred. No. 56;
Conservative 0; Mismatches 25; Indels 0; Gaps 0;

CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGC 70
CCACTAACACTACAGCATGACCCTCTGCTCGCTCGGCTCTTCTGCACGAGAACACTCCC 16

ANCAGT 77
GCCAGT 9

61932 651 bp DNA linear GSS 14-DEC-2000
540D17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ne UUGC1M0540D17 R, genomic survey sequence.
61932
61932.1 GI:11799078
.
musculus (house mouse)
musculus
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
(bases 1 to 651)
n,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
am,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
lly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
derhausern,A. and Wright,D.,Weiss,R.
se whole genome scaffolding with paired end reads from 10kb
smid inserts
ublished (2000)
tact: Robert B. Weiss
versity of Utah Genome Center
versity of Utah
308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
12, USA
: 801 585 5606
: 801 585 7177
il: ddunn@genetics.utah.edu
ert Length: 10000 Std Error: 0.00
e: 0540 row: D column: 17
primer: CACACAGGAAACAGCTATGACC
ss: plasmid ends
quality sequence stop: 651.
Location/Qualifiers
1. .651
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0540D17"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from of pWD42 (gi|4732114|gb|AF129072.1), a copy-n inducible derivative of plasmid R1. The vecto: with adaptors complementary to the insert ada purified. The sheared, adaptored mouse DNA wa: adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Strat: and selected for ampicillin resistance."
ORIGIN

Query Match 37.4%; Score 28.8; DB 28; Length 651;
Best Local Similarity 62.7%; Pred. No. 66;
Matches 42; Conservative 0; Mismatches 25; Indels 0;

Qy 11 CCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNA
Db 420 CCCACTAACACTACAGCATGACCCTCTGCTCGCTCGGCTCTTCTGCACGAGAA

Qy 71 CANCAGT 77
Db 480 TGCCAGT 486

RESULT 15
BF237982/c
LOCUS BF237982 935 bp mRNA linear EST
DEFINITION 601842582F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4(
mRNA sequence.
ACCESSION BF237982
VERSION BF237982.1 GI:11151901
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hc
REFERENCE 1 (bases 1 to 935)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collectio
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LL
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM939 row: d column: 21
High quality sequence stop: 629.
FEATURES Location/Qualifiers
source 1. .935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4079972"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: X
EcoRI; cDNA made by oligo-dT priming. Directio
into EcoRI/XhoI sites using the following 5' a
GGCACGAG(G). Size-selected >500bp for average
1.8kb. Library constructed by Ling Hong in the
of Gerald M. Rubin (University of California,
using ZAP-cDNA synthesis kit (Stratagene) and
II RT (Life Technologies). Note: this is a NIH
Library."
ORIGIN

Query Match 37.1%; Score 28.6; DB 10; Length 935;
Best Local Similarity 65.6%; Pred. No. 90;
Matches 40; Conservative 0; Mismatches 21; Indels 0;

```
ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
ACGTGCCCAGGCTGGTTCCACTCCAGCGTGAGGCTCTGAATGGACTTGTTCTTTGGAGGA 561

3 61

3 560
```

```
749935                  400 bp    mRNA    linear   EST 09-DEC-2003
749935 XGC-gastrula Silurana tropicalis cDNA clone TGas073m07 3',
NA sequence.
749935
749935.1  GI:39640003
.
urana tropicalis (western clawed frog)
urana tropicalis
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
hibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
opodinae; Silurana.
 (bases 1 to 400)
ning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
ger Xenopus tropicalis EST project 2001 (11_2003)
ublished (2003)
tact: Croning MDR
ger Institute
xton, Cambridgeshire, CB10 1SA, UK
il: trop@sanger.ac.uk
ger Xenopus tropicalis EST project 2001
PICALIS_SEQUENCE_ID: TGas073m07.q1kT7
uencing primer: T7
s sequence is from a Xenopus Gene Collection (XGC) library
structed by Aaron M. Zorn.
NA was oligo dT primed from 5ug of poly A+ RNA from stages 10-13
trulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
RI at the 5' end and NotI at the 3' end.
tor: pCS107; Site_1: EcoRI; Site_2: NotI
t: Escherichia coli XL1- blue.
      Location/Qualifiers
      1. .400
      /organism="Silurana tropicalis"
      /mol_type="mRNA"
      /db_xref="taxon:8364"
      /clone="TGas073m07"
      /dev_stage="gastrula (stages 10.5-12 mixed)"
      /lab_host="Escherichia coli XL1-blue"
      /clone_lib="XGC-gastrula"
      /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
      was oligo dT primed from 5ug of poly A+ RNA from stages
      10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
      into pCS107 with EcoRI at the 5' end and NotI at the 3'
      end."
```

```
          36.6%;  Score 28.2;  DB 13;  Length 400;
milarity   65.0%;  Pred. No. 84;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNA 63
ACCCATACTAGCACCATTCTAGCCACCGTGCCCTTATGGTCACATCATTTTGTGGTGGA 96
```

```
58637                   617 bp    DNA     linear   GSS 14-DEC-2000
535I05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ne UUGC1M0535I05 R, genomic survey sequence.
58637
58637.1  GI:11795783
```

```
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Ha
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads fr
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 203(
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0535  row: I  column: 05
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 617.
FEATURES             Location/Qualifiers
     source          1. .617
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0535I05"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resist
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA f
                     musculus C57BL/6J (male) was obtained  from th
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares
                     was hydrodynamically sheared by repeated passa
                     0.005 inch orifice at constant velocity. The s
                     was blunt end-repaired with T4 DNA polymerase
                     polynucleotide kinase. Adaptor oligonucleotide
                     ligated to the blunt ends in high  molar exces
                     adaptored DNA was purified and size-selected f
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-nu
                     inducible derivative of plasmid R1. The vector
                     with adaptors complementary to the insert adap
                     purified. The sheared, adaptored mouse DNA was
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Strata
                     and selected for ampicillin resistance."
ORIGIN
```

```
  Query Match             36.4%;  Score 28;  DB 28;  Length 617;
  Best Local Similarity   60.6%;  Pred. No. 1.2e+02;
  Matches   43;  Conservative    0;  Mismatches   28;  Indels    0;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATT
              |||| | |   ||||||    |   ||| ||||   | |   || |||||||
Db        389 ACTTGCTCTCACCAGCACGGTGAAAGCTTCATTTGCAGAGCGGGCAGGTCAAG

Qy         61 GNACAACAGCC 71
              | |||  ||||
Db        449 GTACAGAAGCC 459


RESULT 18
CE510405/c
LOCUS       CE510405                625 bp    DNA     linear   GSS
```

r-gss-dog-17000327381038 Dog Library Canis familiaris genomic,
omic survey sequence.
10405
10405.1 GI:36827186
.
is familiaris (dog)
is familiaris
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
(bases 1 to 625)
kness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
ch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
ter,J.C.
dog genome: survey sequencing and comparative analysis
ence 301 (5641), 1898-1903 (2003)
75432
12627
tact: Kirkness EF
Institute for Genomic Research
artment of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
kville, MD 20850, USA
: 301-838-0200
: 301-838-0208
il: ekirknes@tigr.org
ss: shotgun.
Location/Qualifiers
1. .625
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site_1: BstXI; Libraries were prepared from
peripheral blood"

36.4%; Score 28; DB 29; Length 625;
milarity 60.6%; Pred. No. 1.2e+02;
Conservative 0; Mismatches 28; Indels 0; Gaps 0;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
| ||| || | ||| | || | |||| |||| || |||| ||||||||
CATACCCATCTCAGAATTTTTTAGGGATCAAATTCTGTATTTTCAGAATATTTGGAGAT 96

NACAACAGCC 71
||| ||
TTCAAATTCC 85

65668 939 bp mRNA linear EST 29-NOV-2002
368312F1 CSEQRBN19 Gallus gallus cDNA clone ChEST271g3 5', mRNA
uence.
65668
65668.1 GI:25955129
.
lus gallus (chicken)
lus gallus
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
hosauria; Aves; Neognathae; Galliformes; Phasianidae;
sianinae; Gallus.
(bases 1 to 939)
rdman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
g,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
omprehensive Collection of Chicken cDNAs
r. Biol. 12 (22), 1965-1969 (2002)
35534
45392
tact: Simon Hubbard
artment of Biomolecular Sciences
versity of Manchester Institute of Science and Technology
IST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
FEATURES Location/Qualifiers
source 1. .939
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="ChEST271g3"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN19"
/note="Organ: ovary; Vector: pBluescript II KS
EcoRI; Site_2: NotI; This normalized library w
constructed from 1 million independent clones.
synthesis was initiated using an oligo(dT) pri
methylated C in the first strand synthesis rea
Following this first strand reaction, double-s
was blunted, ligated to NotI adapters, digeste
EcoRI, size-selected, and cloned into the NotI
compatible sites of a custom modified MCS of t
pBluescript (KS+) vector. The library was norm
rounds using conditions adapted from Soares et
(1994) 91: 9228-9232 and Bonaldo et al., Genom
(1996): 791, except that a significantly longe
reannealing hybridization was used."
ORIGIN

Query Match 36.4%; Score 28; DB 13; Length 939;
Best Local Similarity 63.5%; Pred. No. 1.4e+02;
Matches 40; Conservative 0; Mismatches 23; Indels 0;

Qy 12 CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNAC
|| || | || | | |||| ||||| | | | | | |||||||| | ||
Db 901 CCCGCTTCTTTTTGGGATCACCCTCTTCCCAGGCCGCTTTTTTGGAGGGGTAC

Qy 72 ANC 74
| |
Db 841 ATC 839

RESULT 20
AZ135044/c
LOCUS AZ135044 544 bp DNA linear GSS
DEFINITION OSJNBb0114J07r CUGI Rice BAC Library (EcoRI) Oryza sati
cultivar-group) genomic clone OSJNBb0114J07r, genomic s
sequence.
ACCESSION AZ135044
VERSION AZ135044.1 GI:8214286
KEYWORDS GSS.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poace
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 544)
AUTHORS Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Gen
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 40
High quality sequence stop: 496.

```
      Location/Qualifiers
      1. .544
      /organism="Oryza sativa (japonica cultivar-group)"
      /mol_type="genomic DNA"
      /strain="Japonica"
      /cultivar="Nipponbare"
      /db_xref="taxon:39947"
      /clone="OSJNBb0114J07r"
      /tissue_type="Leaf"
      /lab_host="E. coli DH10B"
      /clone_lib="CUGI Rice BAC Library (EcoRI)"
      /note="Vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
      Rice is the most important food crop in the world. Half of
      the world population, especially those inhabiting highly
      populated areas of the humid tropics and subtropics, rely
      on rice as their primary source of carbohydrate.
      Monocotyledonous rice is a diploid plant (2n=24) with a
      haploid genome equivalent of 431 Mbp (Arumuganathan and
      Earle, 1991). The relatively small genome of rice, three
      times larger than that of Arabidopsis, makes it suitable
      for genomic studies. In order to facilitate positional
      cloning, physical mapping and genome sequencing of rice,
      we have constructed a BAC library from Oryza sativa,
      Nipponbare variety using EcoRI as the cloning enzyme. The
      library contains 55,296 clones with an average insert size
      of 121 Kb providing approximatley 15 haploid genome
      equivalents. The deep coverage allows the isolation a
      particular sequence with a probability of 99.9 %. Three
      high density filters, each containing 18,432 clones
      (doubly spotted), represent the whole library for colony
      screening and can be requested from the Clemson University
      BAC/EST Resource Center (www.genome.clemson.edu)."

          35.8%;  Score 27.6;  DB 28;  Length 544;
milarity  65.0%;  Pred. No. 1.5e+02;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAG 61
||     || | |||   || || | |||| ||||| ||| ||||| ||  ||| | ||
TTGGAGTGCACTGCATTCTTATAACGTCATACTCTGTATGATCAGGGCACATGGTGCAG 208

60588                  781 bp    DNA     linear   GSS 03-NOV-1999
b0015K14r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica
tivar-group) genomic clone nbeb0015K14r, genomic survey
uence.
60588
60588.1  GI:6211045
.
za sativa (japonica cultivar-group)
za sativa (japonica cultivar-group)
aryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
rmatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
hartoideae; Oryzeae; Oryza.
(bases 1 to 781)
g,R.A. and Dean,R.A.
AC End Sequencing Framework to Sequence the Rice Genome
ublished (1998)
tact: Wing RA
mson University Genomics Institute
mson University
) Jordan Hall, Clemson, SC 29634, USA
.: 864 656 7288
:: 864 656 4293
il: rwing@clemson.edu
[ primer: GGAAACAGCTATGACCATG
ss: BAC ends
jh quality sequence start: 37
jh quality sequence stop: 408.
      Location/Qualifiers
```

```
      source          1. .781
                      /organism="Oryza sativa (japonica cultivar-gro
                      /mol_type="genomic DNA"
                      /strain="Japonica"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="nbeb0015K14r"
                      /tissue_type="Leaf"
                      /lab_host="E. coli DH10B"
                      /clone_lib="CUGI Rice BAC Library (EcoRI)"
                      /note="Vector: pBACIndigo; Site_1: EcoRI; Site_
                      Rice is the most important food crop in the wo
                      the world population, especially those inhabit
                      populated areas of the humid tropics and subtr
                      on rice as their primary source of carbohydrat
                      Monocotyledonous rice is a diploid plant (2n=2
                      haploid genome equivalent of 431 Mbp (Arumugan
                      Earle, 1991). The relatively small genome of r
                      times larger than that of Arabidopsis, makes i
                      for genomic studies. In order to facilitate po
                      cloning, physical mapping and genome sequencin
                      we have constructed a BAC library from Oryza s
                      Nipponbare variety using EcoRI as the cloning
                      library contains 55,296 clones with an average
                      of 121 Kb providing approximatley 15 haploid g
                      equivalents. The deep coverage allows the isol
                      particular sequence with a probability of 99.9
                      high density filters, each containing 18,432 c
                      (doubly spotted), represent the whole library
                      screening and can be requested from the Clemso
                      BAC/EST Resource Center (www.genome.clemson.ed
ORIGIN

  Query Match           35.8%;  Score 27.6;  DB 28;  Length 781;
  Best Local Similarity 65.0%;  Pred. No. 1.8e+02;
  Matches   39;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          2 CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTT
              |||     || | |||   || || | |||| ||||| ||| ||||| ||  |
Db        326 CTTGGAGTGCACTGCATTCTTATAACGTCATACTCTGTATGATCAGGGCACAT

RESULT 22
BB588106/c
LOCUS       BB588106                 221 bp    mRNA    linear   EST
DEFINITION  BB588106 RIKEN full-length enriched, 6 days neonate ski
            musculus cDNA clone A030005I02 5', mRNA sequence.
ACCESSION   BB588106
VERSION     BB588106.1  GI:11484650
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
REFERENCE   1  (bases 1 to 221)
  AUTHORS   Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
            Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hiroz
            Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., K
            Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazak
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,
            Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogab
            Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
            Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yos
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-
            Tel: 81-45-503-9222
```

: 81-45-503-9216
il: genome-res@gsc.riken.go.jp,
:http://genome.gsc.riken.go.jp/
ninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
aki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
rmostabilization and thermoactivation of thermolabile enzymes by
halose and its application for the synthesis of full length
A. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
oh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
aru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
zaki,Y. and Hayashizaki,Y.
tomated filtration-based high-throughput plasmid preparation
tem. Genome Res. 9 (5), 463-470 (1999)
rninci,P. and Hayashizaki,Y.
gh-efficiency full-length cDNA cloning. Methods Enzymol. 303,
44 (1999)
ease visit our web site (http://genome.rtc.riken.go.jp) for
ther details.
Location/Qualifiers
1. .221
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A030005I02"
/tissue_type="skin"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 6 days neonate skin"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. "

35.6%; Score 27.4; DB 10; Length 221;
nilarity 59.7%; Pred. No. 1.2e+02;
Conservative 0; Mismatches 29; Indels 0; Gaps 0;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
|| | | || | || | |||||| || | | |||| || ||||||| |
CTAATTCTGCAGATATACACTTTAGCATGATTTTATACATGTCAAGCACATTTGGGAGT 147

NACAACAGCCA 72
|| | || ||
GACTAGAGACA 135

83092 301 bp mRNA linear EST 03-APR-2002
83092 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica tivar-group) cDNA clone E41124, mRNA sequence.
83092
83092.1 GI:13896756
.
za sativa (japonica cultivar-group)
za sativa (japonica cultivar-group)
aryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; rmatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 301)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from panicle (between 3cm to 10cm) (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affr
PROJECT = 'RGP'.
FEATURES Location/Qualifiers
source 1. .301
/organism="Oryza sativa (japonica cultivar-grc
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E41124"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)
ORIGIN

Query Match 35.6%; Score 27.4; DB 9; Length 301;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 40; Conservative 0; Mismatches 24; Indels 0;

Qy 12 CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNAC
||| | ||| || ||||| | | || |||| || || | || ||
Db 241 CCAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAAC

Qy 72 ANCA 75
| ||
Db 181 AACA 178

RESULT 24
CE670554
LOCUS CE670554 304 bp DNA linear GSS
DEFINITION tigr-gss-dog-17000313750722 Dog Library Canis familiari genomic survey sequence.
ACCESSION CE670554
VERSION CE670554.1 GI:36989554
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Can
REFERENCE 1 (bases 1 to 304)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remingt
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M.
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analy
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical C
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .304
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"

/note="Site_1: BstXI; Libraries were prepared from peripheral blood"

35.6%; Score 27.4; DB 29; Length 304;
milarity 73.9%; Pred. No. 1.4e+02;
Conservative 0; Mismatches 12; Indels 0; Gaps 0;

TCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAA 66
|| | |||||||| | |||||| || | ||| ||| || ||||
TCATGAATCATCCTTTCCATGGTAAGCTTCTTTAGAGAAGAACAA 243

82930 330 bp mRNA linear EST 03-APR-2002
82930 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica
tivar-group) cDNA clone E40377, mRNA sequence.
82930
82930.1 GI:13896594
.
za sativa (japonica cultivar-group)
za sativa (japonica cultivar-group)
aryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
rmatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
hartoideae; Oryzeae; Oryza.
(bases 1 to 330)
aki,T. and Yamamoto,K.
e cDNA from panicle (between 3cm to 10cm) (2001)
ublished (2001)
ntact: Takuji Sasaki
ional Institute of Agrobiological Resources
e Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
-8602, Japan
.: 81-298-38-7441
:: 81-298-38-7468
il: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
JECT = 'RGP'.
Location/Qualifiers
1. .330
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E40377"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)"

35.6%; Score 27.4; DB 9; Length 330;
milarity 62.5%; Pred. No. 1.4e+02;
Conservative 0; Mismatches 24; Indels 0; Gaps 0;

CAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGCC 71
||| | ||| || ||||| | | || |||| || || | || |||||||
CAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAACAACAGGC 211

ANCA 75
| ||
AACA 207

183801 447 bp mRNA linear EST 03-APR-2002
183801 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica
ltivar-group) cDNA clone E40741, mRNA sequence.
183801
183801.1 GI:14191590
T.
yza sativa (japonica cultivar-group)
yza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poace
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 447)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from panicle (between 3cm to 10cm) (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba,
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affr
PROJECT = 'RGP'.
FEATURES Location/Qualifiers
source 1. .447
/organism="Oryza sativa (japonica cultivar-grc
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E40741"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)
ORIGIN

Query Match 35.6%; Score 27.4; DB 9; Length 447;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 40; Conservative 0; Mismatches 24; Indels 0;

Qy 12 CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNAC
||| | ||| || ||||| | | || |||| || || | || ||
Db 388 CCAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAAC

Qy 72 ANCA 75
| |
Db 328 AACA 325

RESULT 27
AU223051/c
LOCUS AU223051 451 bp mRNA linear EST
DEFINITION AU223051 Rice shoot Oryza sativa (japonica cultivar-grc
clone S3117, mRNA sequence.
ACCESSION AU223051
VERSION AU223051.1 GI:15008663
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poace
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 451)
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from etiolated shoot (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba,
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affr
PROJECT = 'RGP'.
FEATURES Location/Qualifiers
source 1. .451
/organism="Oryza sativa (japonica cultivar-grc
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="S3117"
/dev_stage="Etiolated shoot (8 days old)"

```
       /clone_lib="Rice shoot"
       /note="Etiolated shoot (8 days old)"


           35.6%;  Score 27.4;  DB 9;  Length 451;
imilarity   62.5%;  Pred. No. 1.6e+02;
;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGCC 71
||| |  |||  || |||||   | |   ||   |||| || || |  || |||||||
CCAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAACAACAGGC 185

ANCA 75
| ||
AACA 181



222931                 453 bp    mRNA    linear   EST 03-APR-2002
222931 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
one S1940, mRNA sequence.
222931
222931.1  GI:15008543
T.
yza sativa (japonica cultivar-group)
yza sativa (japonica cultivar-group)
karyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
rhartoideae; Oryzeae; Oryza.
 (bases 1 to 453)
saki,T. and Yamamoto,K.
ce cDNA from etiolated shoot (2001)
published (2001)
ntact: Takuji Sasaki
tional Institute of Agrobiological Resources
ce Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
5-8602, Japan
l: 81-298-38-7441
x: 81-298-38-7468
ail: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
OJECT = 'RGP'.
       Location/Qualifiers
       1. .453
       /organism="Oryza sativa (japonica cultivar-group)"
       /mol_type="mRNA"
       /cultivar="Nipponbare"
       /db_xref="taxon:39947"
       /clone="S1940"
       /dev_stage="Etiolated shoot (8 days old)"
       /clone_lib="Rice shoot"
       /note="Etiolated shoot (8 days old)"


           35.6%;  Score 27.4;  DB 9;  Length 453;
imilarity   62.5%;  Pred. No. 1.6e+02;
;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGCC 71
||| |  |||  || |||||   | |   ||   |||| || || |  || |||||||
CCAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAACAACAGGC 342

ANCA 75
| ||
AACA 338



183799                 454 bp    mRNA    linear   EST 03-APR-2002
183799 Rice panicle (between 3cm to 10cm) Oryza sativa (japonica
ltivar-group) cDNA clone E40701, mRNA sequence.
183799
VERSION     AU183799.1  GI:14191588
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from panicle (between 3cm to 10cm) (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.aff
            PROJECT = 'RGP'.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Oryza sativa (japonica cultivar-gr
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="E40701"
                     /tissue_type="panicle"
                     /dev_stage="panicle (between 3cm to 10cm)"
                     /clone_lib="Rice panicle (between 3cm to 10cm
ORIGIN

  Query Match           35.6%;  Score 27.4;  DB 9;  Length 454;
  Best Local Similarity   62.5%;  Pred. No. 1.6e+02;
  Matches   40;  Conservative    0;  Mismatches   24;  Indels    0

Qy         12 CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNA
              ||| |  |||  || |||||   | |   ||   |||| || || |  || |
Db        395 CCAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAA

Qy         72 ANCA 75
              | ||
Db        335 AACA 332


RESULT 30
AU172933/c
LOCUS       AU172933                470 bp    mRNA    linear   ES
DEFINITION  AU172933 Rice panicle (between 3cm to 10cm) Oryza sati
            cultivar-group) cDNA clone E41160, mRNA sequence.
ACCESSION   AU172933
VERSION     AU172933.1  GI:12622720
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from panicle (between 3cm to 10cm) (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Takuji Sasaki
            Naticnal Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.aff
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Oryza sativa (japonica cultivar-gr
                     /mol_type="mRNA"
```

/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E41160"
/tissue_type="panicle"
/dev_stage="panicle (between 3cm to 10cm)"
/clone_lib="Rice panicle (between 3cm to 10cm)"

35.6%; Score 27.4; DB 9; Length 470;
milarity 62.5%; Pred. No. 1.7e+02;
Conservative 0; Mismatches 24; Indels 0; Gaps 0;

:CAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGCC 71
:CAACCACATAGTATCATCAGAGTTTTTCTGCATAGGTAATCTGAACAAGAACAACAGGC 352

ANCA 75
AACA 348

487714 482 bp mRNA linear EST 07-FEB-2002
n2n.pk005.g10 Normalized Chicken Breast Muscle, Leg Muscle, and
iphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
one pgm2n.pk005.g10 5' similar to dbj|BAB26278.1 (AK009426)
tative [Mus musculus], mRNA sequence.
487714
487714.1 GI:18608644
T.
llus gallus (chicken)
llus gallus
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
chosauria; Aves; Neognathae; Galliformes; Phasianidae;
asianinae; Gallus.
(bases 1 to 482)
gburn,L.A. and Monsonego-Ornan,E.
Ts from Normalized Chicken Breast Muscle, Leg Muscle, and
iphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
oject
published (2002)
ntact: Larry A. Cogburn
iversity of Delaware
wnsend Hall, Newark, DE 19717, USA
l: 302-831-1335
x: 302-831-2822
ail: cogburn@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
1. .482
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk005.g10"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
,11 weeks);growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMDH10B"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/note="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

35.6%; Score 27.4; DB 12; Length 482;
Similarity 62.5%; Pred. No. 1.7e+02;

Matches 40; Conservative 0; Mismatches 24; Indels 0;

Qy 6 CNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAG
Db 282 CTTAGCCCTGCACCAGTCCAGCAACATCTGCTTAATGCTGTTGGCATTGGGGA

Qy 66 ACAG 69
Db 222 GCTG 219

RESULT 32
CB685142
LOCUS CB685142 687 bp mRNA linear EST
DEFINITION OSJNEf15F03.r OSJNEf Oryza sativa (japonica cultivar-g
clone OSJNEf15F03 3', mRNA sequence.
ACCESSION CB685142
VERSION CB685142.1 GI:29688867
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; T
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poac
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 687)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the int
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucso
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: F column: 03
Seq primer: gga aac agc tat gac cat g.
FEATURES Location/Qualifiers
source 1. .687
/organism="Oryza sativa (japonica cultivar-g
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEf15F03"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEf"
/note="Vector: pBluescript II KS +; Site_1:
XhoI; Uninfected Control"
ORIGIN

Query Match 35.6%; Score 27.4; DB 14; Length 687
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 40; Conservative 0; Mismatches 24; Indels

Qy 12 CCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGN
Db 25 CCAACCACATAGTATCATCCTAGTTTTTCTGCATAGGTAATCTGAACAAGA

Qy 72 ANCA 75
Db 85 AACA 88

RESULT 33
CE474611

```
:474611                 782 bp    DNA     linear   GSS 27-SEP-2003
:gr-gss-dog-17000364151175 Dog Library Canis familiaris genomic,
:nomic survey sequence.
:474611
:474611.1  GI:36778061
:S.
:nis familiaris (dog)
:nis familiaris
:karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
:mmalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
  (bases 1 to 782)
:rkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
:sch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
:nter,J.C.
:e dog genome: survey sequencing and comparative analysis
:ience 301 (5641), 1898-1903 (2003)
:875432
512627
:ntact: Kirkness EF
:e Institute for Genomic Research
:partment of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
:ckville, MD 20850, USA
:l: 301-838-0200
x: 301-838-0208
:ail: ekirknes@tigr.org
ass: shotgun.
       Location/Qualifiers
       1. .782
       /organism="Canis familiaris"
       /mol_type="genomic DNA"
       /strain="Standard Poodle"
       /db_xref="taxon:9615"
       /clone_lib="Dog Library"
       /note="Site_1: BstXI; Libraries were prepared from
       peripheral blood"


           35.6%;  Score 27.4;  DB 29;  Length 782;
imilarity   62.5%;  Pred. No. 2.1e+02;
;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

CCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNACAACAGC 70
| | |   || | | || ||  |||||||| || ||| |    | |||||  ||| ||||
CTCTGGCACAGTGTGGCCTCCCCCTCTGCAGGGCCAGCTTGCCTAGAGGACCACACCAGC 605

CANC 74
|  |
CTCC 609



190117                  984 bp    mRNA    linear   EST 04-SEP-2002
ENCOURT_8076683 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6089012
, mRNA sequence.
190117
190117.1  GI:22704101
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  (bases 1 to 984)
H-MGC http://mgc.nci.nih.gov/.
tional Institutes of Health, Mammalian Gene Collection (MGC)
published (1999)
ntact: Robert Strausberg, Ph.D.
ail: cgapbs-r@mail.nih.gov
ssue Procurement: DCTD/DTP
DNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
NA Sequencing by: Agencourt Bioscience Corporation
lone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM2328  row: b  column: 21
          High quality sequence stop: 661.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6089012"
                     /tissue_type="melanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_112"
                     /note="Organ: skin; Vector: pOTB7; Site_1: Xh
                     EcoRI; cDNA made by oligo-dT priming. Directi
                     into EcoRI/XhoI sites using the following 5'
                     GGCACGAG(G). Library constructed  by Ling Hor
                     laboratory of Gerald M. Rubin (University of
                     Berkeley) using ZAP-cDNA synthesis kit (Strat
                     Superscript II RT (Life Technologies). Note:
                     NIH_MGC Library."
ORIGIN

  Query Match           35.6%;  Score 27.4;  DB 13;  Length 984;
  Best Local Similarity   65.8%;  Pred. No. 2.3e+02;
  Matches   52;  Conservative    0;  Mismatches   25;  Indels    2

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG--GTCAGGTC
              || |||| || ||||||| ||||||         |   | | |||     ||
Db        791 ACTTACTCACCCCAGCATCATTCTTCTCCTCTTGGCAGAATGTCTGTAGCCC

Qy         59 GAGNACAACAGCCANCAGT 77
              ||  ||| |||||| ||||
Db        851 GAGTACACCAGCCATCAGT 869


RESULT 35
BM557397
LOCUS       BM557397                1068 bp    mRNA    linear   ES
DEFINITION  AGENCOURT_6561703 NIH_MGC_72 Homo sapiens cDNA clone I
            5', mRNA sequence.
ACCESSION   BM557397
VERSION     BM557397.1  GI:18799358
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collecti
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (I
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution informatic
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12255  row: n  column: 11
            High quality sequence stop: 643.
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5547850"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site
```

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

35.6%; Score 27.4; DB 12; Length 1068;
imilarity 65.8%; Pred. No. 2.4e+02;
; Conservative 0; Mismatches 25; Indels 2; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG--GTCAGGTCATTTGGAG 58
|||||| |||||||||| |||||| | | | || | |||||||
ACTTACTCAGCCCAGCATCATTCTTCTTCTCTTGGCAGATTGTCTGTAACCGATTTGGAG 907

GAGNACAACAGCCANCAGT 77
||| ||||||| || ||||
GAGTACAACAGGCATCAGT 926

595860 462 bp mRNA linear EST 25-FEB-2002
000687477444 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
600449734692 5', mRNA sequence.
595860
595860.1 GI:18891725
T.
opheles gambiae (African malaria mosquito)
opheles gambiae
karyota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
optera; Endopterygota; Diptera; Nematocera; Culicoidea;
opheles.
(bases 1 to 462)
lt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
arlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
lera Anopheles gambiae EST project
ıpublished (2002)
ntact: Holt R.A.
lera Genomics
; w. Gude Dr., Rockville, MD 20850, USA
el: 2404533151
ax: 2404534580
ıail: HoltRA@celera.com
late: NU010049UW row: N column: 22
eq primer: M13 Reverse.
Location/Qualifiers
1. .462
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449734692"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)"

35.3%; Score 27.2; DB 12; Length 462;
Similarity 57.9%; Pred. No. 1.9e+02;
4; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
| || | |||| | | ||| || | | ||||| |||||| ||| | || |||
AGTTTCGCAGCTGATCCCTTTTCCGCCAGTAGCTTCTGCGTGGTCAAATCAATCGGGGGA 277

GNACAACAGCCANCAG 76
| | || | ||

Db 276 GCAACTTCGCGACGAG 261

RESULT 37
BM644507/c
LOCUS BM644507 465 bp mRNA linear EST
DEFINITION 17000687314460 A.Gam.ad.cDNA1 Anopheles gambiae cDNA c
19600449634703 5', mRNA sequence.
ACCESSION BM644507
VERSION BM644507.1 GI:18944018
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pte
Neoptera; Endopterygota; Diptera; Nematocera; Culicoid
Anopheles.
REFERENCE 1 (bases 1 to 465)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C
Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 w. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004ABV row: I column: 05
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .465
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin
chromosome)"
/db_xref="taxon:7165"
/clone="19600449634703"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cDNA1"
/note="Vector: pSport1; Site_1: SalI; Site_2
adult mosquitoes (mixed sex) frozen on liqui
cDNA inserts >500 bp cloned directionally i
Not 1 site is 3'. Clones available through
Research and Reference Reagent Resource Cen
(www.malaria.mr4.org)."
ORIGIN

Query Match 35.3%; Score 27.2; DB 12; Length 465
Best Local Similarity 57.9%; Pred. No. 1.9e+02;
Matches 44; Conservative 0; Mismatches 32; Indels

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCA
| || | |||| | | ||| || | | ||||| |||||| |||
Db 306 AGTTTCGCAGCTGATCCCTTTTCCGCCAGTAGCTTCTGCGTGGTCAAATCA

Qy 61 GNACAACAGCCANCAG 76
| | || | ||
Db 246 GCAACTTCGCGACGAG 231

RESULT 38
BM648778/c
LOCUS BM648778 481 bp mRNA linear E
DEFINITION 17000687367686 A.Gam.ad.cDNA1 Anopheles gambiae cDNA
19600449621320 5', mRNA sequence.
ACCESSION BM648778
VERSION BM648778.1 GI:18948289
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pt
Neoptera; Endopterygota; Diptera; Nematocera; Culicoi

```
pheles.
(bases 1 to 481)
t,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
rlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
era Anopheles gambiae EST project
ublished (2002)
tact: Holt R.A.
era Genomics
w. Gude Dr., Rockville, MD 20850, USA
: 2404533151
: 2404534580
il: HoltRA@celera.com
te: NU01004N8I  row: K  column: 14
 primer: M13 Reverse.
      Location/Qualifiers
      1. .481
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="RSP-ST (Reduced susc. to Permethrin - std.
      chromosome)"
      /db_xref="taxon:7165"
      /clone="19600449621320"
      /dev_stage="Adult"
      /lab_host="DH10b"
      /clone_lib="A.Gam.ad.cDNA1"
      /note="Vector: pSport1; Site_1: Sal1; Site_2: Not1; Whole
      adult mosquitoes (mixed sex) frozen on liquid nitrogen.
      cDNA  inserts >500 bp cloned directionally into pSport 1.
      Not 1 site is  3'. Clones available through the Malaria
      Research and Reference  Reagent Resource Center
      (www.malaria.mr4.org)."


          35.3%;  Score 27.2;  DB 12;  Length 481;
milarity   57.9%;  Pred. No. 2e+02;
  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
 || | |||| | |    |||   ||  | | ||||| ||||||  ||| | || |||
GTTTCGCAGCTGATCCCTTTTCCGCCAGTAGCTTCTGCGTGGTCAAATCAATCGGGGGA 293

NACAACAGCCANCAG 76
  |     || |  ||
CAACTTCGCGACGAG 277



08F9W                  482 bp    mRNA    linear   HTC 07-JAN-2003
gle read from an extremity of a full-length cDNA clone made from
pheles gambiae total adult females. 5-PRIME end of clone
AAA15AA04 of strain 6-9 of Anopheles gambiae (African malaria
quito).
09216
09216.1  GI:27558436
'.
pheles gambiae (African malaria mosquito)
pheles gambiae
aryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
ptera; Endopterygota; Diptera; Nematocera; Culicoidea;
pheles.
(bases 1 to 482)
ioscope.
ect Submission
mitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
eb : www.genoscope.cns.fr)
      Location/Qualifiers
      1. .482
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="6-9"
      /db_xref="taxon:7165"
                  /clone="FK0AAA15AA04"
                  /plasmid="pME18S-FL"
                  /note="end : 5-PRIME"
ORIGIN

  Query Match           35.3%;  Score 27.2;  DB 11;  Length 482;
  Best Local Similarity   57.9%;  Pred. No. 2e+02;
  Matches   44;  Conservative    0;  Mismatches   32;  Indels    0;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATT
              | || | |||| | |    |||   ||  | | ||||| ||||||  ||| |
Db        280 AGTTTCGCAGCTGATCCCTTTTCCGCCAGTAGCTTCTGCGTGGTCAAATCAAT

Qy         61 GNACAACAGCCANCAG 76
              | |     || |  ||
Db        220 GCAACTTCGCGACGAG 205


RESULT 40
BM643718/c
LOCUS       BM643718                 487 bp    mRNA    linear   EST
DEFINITION  17000687312981 A.Gam.ad.cDNA1 Anopheles gambiae cDNA cl
            19600449660604 5', mRNA sequence.
ACCESSION   BM643718
VERSION     BM643718.1  GI:18943229
KEYWORDS    EST.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pter
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoide
            Anopheles.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.
            Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
  TITLE     Celera Anopheles gambiae EST project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Holt R.A.
            Celera Genomics
            45 w. Gude Dr., Rockville, MD 20850, USA
            Tel: 2404533151
            Fax: 2404534580
            Email: HoltRA@celera.com
            Plate: NU01004HNP  row: P  column: 10
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Anopheles gambiae"
                     /mol_type="mRNA"
                     /strain="RSP-ST (Reduced susc. to Permethrin -
                     chromosome)"
                     /db_xref="taxon:7165"
                     /clone="19600449660604"
                     /dev_stage="Adult"
                     /lab_host="DH10b"
                     /clone_lib="A.Gam.ad.cDNA1"
                     /note="Vector: pSport1; Site_1: Sal1; Site_2:
                     adult mosquitoes (mixed sex) frozen on liquid
                     cDNA  inserts >500 bp cloned directionally int
                     Not 1 site is  3'. Clones available through th
                     Research and Reference  Reagent Resource Cente
                     (www.malaria.mr4.org)."
ORIGIN

  Query Match           35.3%;  Score 27.2;  DB 12;  Length 487;
  Best Local Similarity   57.9%;  Pred. No. 2e+02;
  Matches   44;  Conservative    0;  Mismatches   32;  Indels    0;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATT
              | || | |||| | |    |||   ||  | | ||||| ||||||  ||| |
Db        327 AGTTTCGCAGCTGATCCCTTTTCCGCCAGTAGCTTCTGCGTGGTCAAATCAAT

Qy         61 GNACAACAGCCANCAG 76
              | |     || |  ||
```

ɜCAACTTCGCGACGAG 252

ed: April 13, 2004, 20:14:23
7.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

icleic search, using sw model

April 13, 2004, 19:12:18 ; Search time 174.5 Seconds
(without alignments)
1655.373 Million cell updates/sec

US-10-090-326-25
77
1 acttacncagcccagcanca.....ggagnacaacagccancagt 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

2475585 seqs, 1875730760 residues

hits satisfying chosen parameters: 4951170

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

. is the number of results predicted by chance to have a
ater than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

Query Match	Length	DB	ID	Description
39.0	1059	10	US-09-940-925A-58	Sequence 58, Appl
39.0	1059	10	US-09-940-925A-59	Sequence 59, Appl
39.0	1059	10	US-09-941-193A-58	Sequence 58, Appl
39.0	1059	10	US-09-941-193A-59	Sequence 59, Appl
39.0	1587	10	US-09-940-925A-60	Sequence 60, Appl
39.0	1587	10	US-09-940-925A-61	Sequence 61, Appl
39.0	1587	10	US-09-941-193A-60	Sequence 60, Appl
39.0	1587	10	US-09-941-193A-61	Sequence 61, Appl
39.0	1906	13	US-10-011-436-1	Sequence 1, Appli
39.0	1964	15	US-10-117-937-5	Sequence 5, Appli
39.0	2041	9	US-09-834-975-959	Sequence 959, App
39.0	6408	9	US-09-996-128A-1	Sequence 1, Appli
38.2	918	15	US-10-027-632-30884	Sequence 30884, A
38.2	918	15	US-10-027-632-30885	Sequence 30885, A
36.9	100	14	US-10-057-810-41	Sequence 41, Appl
36.9	100	14	US-10-057-828-41	Seque:
36.9	3075	15	US-10-369-493-24411	Seque:
36.9	3078	15	US-10-410-012-8	Seque:
36.9	3150	14	US-10-162-688-1	Seque:
36.9	3157	9	US-09-965-703-74	Sequen:
36.9	3284	14	US-10-157-305A-186	Seque:
36.9	3284	14	US-10-157-391-186	Seque:
36.9	3284	14	US-10-157-096-186	Seque:
36.9	3284	14	US-10-157-302-186	Seque:
36.9	3284	14	US-10-157-215A-186	Seque:
36.9	3284	14	US-10-157-299-186	Seque:
36.9	3284	14	US-10-154-951B-186	Seque:
36.9	3284	14	US-10-156-831-186	Seque:
36.9	3284	14	US-10-157-147-186	Seque:
36.9	3284	14	US-10-157-166-186	Seque:
36.9	3284	14	US-10-156-902-186	Seque:
36.9	3284	14	US-10-157-318-186	Seque:
36.9	3284	15	US-10-156-811-186	Seque:
36.9	3284	15	US-10-157-320A-186	Seque:
36.9	3284	15	US-10-157-418A-186	Seque:
36.9	3284	15	US-10-157-317-186	Seque:
36.9	3284	15	US-10-157-339-186	Seque:
36.9	3396	14	US-10-338-411-32	Seque:
36.9	3396	15	US-10-389-640-32	Seque:
36.9	4279	9	US-09-956-998A-1	Sequen:
36.9	5646	12	US-10-270-176-39	Seque:
36.9	5646	12	US-10-270-176-41	Seque:
36.9	5706	12	US-10-270-176-40	Seque:
36.9	5739	12	US-10-270-176-9	Seque:
36.9	5750	12	US-10-270-176-32	Seque:

ALIGNMENTS

RESULT 1
US-09-940-925A-58
; Sequence 58, Application US/09940925A
; Publication No. US20030054338A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,925A
; FILING DATE: 10-Jun-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid

```
TRANDEDNESS: double
'OPOLOGY: linear
LE TYPE: DNA (genomic)
ICE DESCRIPTION: SEQ ID NO: 58:
58

            39.0%;  Score 30;  DB 10;  Length 1059;
.milarity   67.9%;  Pred. No. 0.17;
   Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

AGNACAACAGCCANCAGT 77
| ||||||||||| ||||
AGTACAACAGCCATCAGT 552



59
Application US/09940925A
Io. US20030054338A1
'ORMATION:
'ANT: BROW, MARY ANN D.
      LYAMICHEV, VICTOR I.
      OLIVE, DAVID M.
 OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
                PATHOGENS
R OF SEQUENCES: 165
SPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
TER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
)PERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
NT APPLICATION DATA:
APPLICATION NUMBER: US/09/940,925A
FILING DATE: 10-Jun-2002
CLASSIFICATION: <Unknown>
NEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01756
)MMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
N FOR SEQ ID NO: 59:
NCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
JLE TYPE: DNA (genomic)
NCE DESCRIPTION: SEQ ID NO: 59:
-59

            39.0%;  Score 30;  DB 10;  Length 1059;
imilarity   67.9%;  Pred. No. 0.17;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||||  |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

AGNACAACAGCCANCAGT 77
||  ||||||||||  ||||
```

```
Db        535 AGTACAACAGCCATCAGT 552


RESULT 3
US-09-941-193A-58
; Sequence 58, Application US/09941193A
; Publication No. US20030108873A1
;   GENERAL INFORMATION:
;        APPLICANT: BROW, MARY ANN D.
;                   LYAMICHEV, VICTOR I.
;                   OLIVE, DAVID M.
;        TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
;                            PATHOGENS
;        NUMBER OF SEQUENCES: 165
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MEDLEN & CARROLL
;             STREET: 220 MONTGOMERY STREET, SUITE 2200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: UNITED STATES OF AMERICA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/941,193A
;             FILING DATE: 28-Aug-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: CARROLL, PETER G.
;             REGISTRATION NUMBER: 32,837
;             REFERENCE/DOCKET NUMBER: FORS-01756
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 58:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1059 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-941-193A-58

  Query Match             39.0%;  Score 30;  DB 10;  Length 1059;
  Best Local Similarity   67.9%;  Pred. No. 0.17;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              ||  ||||||||||  ||||
Db        535 AGTACAACAGCCATCAGT 552


RESULT 4
US-09-941-193A-59
; Sequence 59, Application US/09941193A
; Publication No. US20030108873A1
;   GENERAL INFORMATION:
;        APPLICANT: BROW, MARY ANN D.
;                   LYAMICHEV, VICTOR I.
;                   OLIVE, DAVID M.
;        TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
;                            PATHOGENS
;        NUMBER OF SEQUENCES: 165
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MEDLEN & CARROLL
```

TREET: 220 MONTGOMERY STREET, SUITE 2200
ITY: SAN FRANCISCO
TATE: CALIFORNIA
OUNTRY: UNITED STATES OF AMERICA
IP: 94104
ER READABLE FORM:
EDIUM TYPE: Floppy disk
OMPUTER: IBM PC compatible
PERATING SYSTEM: PC-DOS/MS-DOS
OFTWARE: PatentIn Release #1.0, Version #1.30
T APPLICATION DATA:
PPLICATION NUMBER: US/09/941,193A
ILING DATE: 28-Aug-2001
LASSIFICATION: <Unknown>
EY/AGENT INFORMATION:
AME: CARROLL, PETER G.
EGISTRATION NUMBER: 32,837
EFERENCE/DOCKET NUMBER: FORS-01756
MMUNICATION INFORMATION:
ELEPHONE: (415) 705-8410
ELEFAX: (415) 397-8338
FOR SEQ ID NO: 59:
CE CHARACTERISTICS:
ENGTH: 1059 base pairs
YPE: nucleic acid
TRANDEDNESS: double
OPOLOGY: linear
LE TYPE: DNA (genomic)
CE DESCRIPTION: SEQ ID NO: 59:
59

39.0%; Score 30; DB 10; Length 1059;
milarity 67.9%; Pred. No. 0.17;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| |||||| | | || | | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 552

60
Application US/09940925A
o. US20030054338A1
ORMATION:
ANT: BROW, MARY ANN D.
LYAMICHEV, VICTOR I.
OLIVE, DAVID M.
OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
PATHOGENS
OF SEQUENCES: 165
PONDENCE ADDRESS:
DDRESSEE: MEDLEN & CARROLL
TREET: 220 MONTGOMERY STREET, SUITE 2200
ITY: SAN FRANCISCO
TATE: CALIFORNIA
OUNTRY: UNITED STATES OF AMERICA
IP: 94104
ER READABLE FORM:
EDIUM TYPE: Floppy disk
OMPUTER: IBM PC compatible
PERATING SYSTEM: PC-DOS/MS-DOS
OFTWARE: PatentIn Release #1.0, Version #1.30
T APPLICATION DATA:
PPLICATION NUMBER: US/09/940,925A
ILING DATE: 10-Jun-2002
LASSIFICATION: <Unknown>
EY/AGENT INFORMATION:
AME: CARROLL, PETER G.

; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-940-925A-60

Query Match 39.0%; Score 30; DB 10; Length 1587;
Best Local Similarity 67.9%; Pred. No. 0.18;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
|||||| |||||||||| |||||| | | || | | | |||
Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 842 AGTACAACAGCCATCAGT 859

RESULT 6
US-09-940-925A-61
; Sequence 61, Application US/09940925A
; Publication No. US20030054338A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,925A
; FILING DATE: 10-Jun-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-940-925A-61

Query Match 39.0%; Score 30; DB 10; Length 1587;

milarity 67.9%; Pred. No. 0.18;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 859

60
Application US/09941193A
o. US20030108873A1
ORMATION:
ANT: BROW, MARY ANN D.
LYAMICHEV, VICTOR I.
OLIVE, DAVID M.
OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
PATHOGENS
OF SEQUENCES: 165
PONDENCE ADDRESS:
DDRESSEE: MEDLEN & CARROLL
TREET: 220 MONTGOMERY STREET, SUITE 2200
ITY: SAN FRANCISCO
TATE: CALIFORNIA
OUNTRY: UNITED STATES OF AMERICA
IP: 94104
ER READABLE FORM:
EDIUM TYPE: Floppy disk
OMPUTER: IBM PC compatible
PERATING SYSTEM: PC-DOS/MS-DOS
OFTWARE: PatentIn Release #1.0, Version #1.30
T APPLICATION DATA:
PPLICATION NUMBER: US/09/941,193A
ILING DATE: 28-Aug-2001
LASSIFICATION: <Unknown>
EY/AGENT INFORMATION:
AME: CARROLL, PETER G.
EGISTRATION NUMBER: 32,837
EFERENCE/DOCKET NUMBER: FORS-01756
MMUNICATION INFORMATION:
ELEPHONE: (415) 705-8410
ELEFAX: (415) 397-8338
FOR SEQ ID NO: 60:
CE CHARACTERISTICS:
ENGTH: 1587 base pairs
YPE: nucleic acid
TRANDEDNESS: double
OPOLOGY: linear
LE TYPE: DNA (genomic)
CE DESCRIPTION: SEQ ID NO: 60:
60

39.0%; Score 30; DB 10; Length 1587;
milarity 67.9%; Pred. No. 0.18;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 859

61
Application US/09941193A
o. US20030108873A1

; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193A
; FILING DATE: 28-Aug-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-941-193A-61

Query Match 39.0%; Score 30; DB 10; Length 1587;
Best Local Similarity 67.9%; Pred. No. 0.18;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
Db 842 AGTACAACAGCCATCAGT 859

RESULT 9
US-10-011-436-1
; Sequence 1, Application US/10011436
; Publication No. US20020128200A1
; GENERAL INFORMATION:
; APPLICANT: Boon-Falleur, Thierry; Brichard, Vincent; Van P
; Aline;
; DePlaen, Etienne; Coulie, Pierre; Renauld, Jean
; Lethe, Bernard
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INDIVIDUALS SUF
; A
; CELLULAR ABNORMALITY SOME OF WHOSE ABN
; LEUKOCYTE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: District of Colombia
; COUNTRY: USA

IP: 20004
ER READABLE FORM:
EDIUM TYPE: Diskette, 3.25inch, 1.44MB storage
OMPUTER: IBM PS/2
PERATING SYSTEM: PC-DOS
OFTWARE: Wordperfect
T APPLICATION DATA:
PPLICATION NUMBER: US/10/011,436
ILING DATE: 11-Dec-2001
LASSIFICATION: 435
APPLICATION DATA:
PPLICATION NUMBER: 08/511,011
ILING DATE: 5-AUGUST-1995
PPLICATION NUMBER: 08/054,714
ILING DATE: 28-APRIL-1993
PPLICATION NUMBER: 07/994,928
ILING DATE: 22-DECEMBER-1992
EY/AGENT INFORMATION:
AME: Mary Anne Schofield
EGISTRATION NUMBER: 36,669
EFERENCE/DOCKET NUMBER: LUD 5299.9 DIV CON
MMUNICATION INFORMATION:
ELEPHONE: (202) 662-0200
ELEFAX: (202) 662-4643
FOR SEQ ID NO: 1:
CE CHARACTERISTICS:
ENGTH: 1906 base pairs
YPE: nucleic acid
TRANDEDNESS: single
OPOLOGY: linear
CE DESCRIPTION: SEQ ID NO: 1:

39.0%; Score 30; DB 13; Length 1906;
milarity 67.9%; Pred. No. 0.19;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||      |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 865

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 883
```

pplication US/10117937
o. US20030220239A1
MATION:
TL IMMUNO THERAPIES CORP.
SIMARD, John, J.L.
DIAMOND, David, C.
LIU, Liping
XIE, Zhidong
ENTION: EPITOPE SEQUENCES
CE: CTLIMM.027A
ICATION NUMBER: US/10/117,937
NG DATE: 2002-04-04
ATION NUMBER: US 60/282,211
DATE: 2001-04-06
ATION NUMBER: US 60/337,017
DATE: 2001-11-07
ATION NUMBER: US 60/363,210
DATE: 2002-03-07
Q ID NOS: 602
stSEQ for Windows Version 4.0

4

omo sapiens

Query Match 39.0%; Score 30; DB 15; Length 1964;
Best Local Similarity 67.9%; Pred. No. 0.19;
Matches 53; Conservative 0; Mismatches 24; Indels 1

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||      |   | |  ||  |   | |
Db        864 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        924 AGTACAACAGCCATCAGT 941
```

RESULT 11
US-09-834-975-959
; Sequence 959, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENT
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 959
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2041)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-959

Query Match 39.0%; Score 30; DB 9; Length 2041;
Best Local Similarity 67.9%; Pred. No. 0.2;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||      |   | |  ||  |   | |
Db        907 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        967 AGTACAACAGCCATCAGT 984
```

RESULT 12
US-09-996-128A-1
; Sequence 1, Application US/09996128A
; Patent No. US20020150589A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Alan
; APPLICANT: Bergman, Phillip
; APPLICANT: Wolchok, Jedd
; TITLE OF INVENTION: Compositions for treatment of Melanoma and M
; FILE REFERENCE: MSK.P-026-3
; CURRENT APPLICATION NUMBER: US/09/996,128A
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US 09/627,694
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 09/308,697
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: PCT/US97/22669

```
G DATE: 1997-12-10
CATION NUMBER: US 60/036,419
G DATE: 1997-02-18
CATION NUMBER: US 60/032,535
G DATE: 1996-12-10
CATION NUMBER: US 60/180,651
G DATE: 2000-01-26
EQ ID NOS: 2
atentIn version 3.0

08

Artificial Sequence

misc_feature
()..()
RMATION: vector containing human tyrosinase
-1

           39.0%;  Score 30;  DB 9;  Length 6408;
imilarity   67.9%;  Pred. No. 0.27;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 5356

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 5374



30884/c
84, Application US/10027632
No. US20030204075A9
RMATION:
Wang, David G.
VENTION: Identification and Mapping of Single Nucleotide
VENTION:  Polymorphisms in the Human Genome
NCE: 108827.129
LICATION NUMBER: US/10/027,632
ING DATE:  2002-04-30
CATION NUMBER: US 60/218,006
G DATE: 2000-07-12
CATION NUMBER: US 60/198,676
G DATE: 2000-04-20
CATION NUMBER: US 60/193,483
G DATE: 2000-03-29
CATION NUMBER: US 60/185,218
G DATE: 2000-02-24
CATION NUMBER: US 60/167,363
G DATE: 1999-11-23
CATION NUMBER: US 60/156,358
G DATE: 1999-09-28
CATION NUMBER: US 60/146,002
G DATE: 1999-08-09
EQ ID NOS: 325720
astSEQ for Windows Version 4.0
884
8

Human
30884

           38.2%;  Score 29.4;  DB 15;  Length 918;
imilarity   59.2%;  Pred. No. 0.27;
;  Conservative    2;  Mismatches   27;  Indels    0;  Gaps    0;

ACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGNAC 64
||  |||||  ||   ||| |  | :  || |   | | |  | | || ||  ||||| ||
ACACAGCCAGGCTAGATTATCTCGYTGTCATAATCTTTGCAAAGGCAGTTTCAGGAGAAC 248

Qy          65 AACAGCCANCA 75
               ||:|||||  ||
Db         247 AAYAGCCACCA 237


RESULT 14
US-10-027-632-30885/c
; Sequence 30885, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nuclec
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30885
;   LENGTH: 918
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-30885

  Query Match             38.2%;  Score 29.4;  DB 15;  Length 918;
  Best Local Similarity   59.2%;  Pred. No. 0.27;
  Matches   42;  Conservative    2;  Mismatches   27;  Indels    0

Qy           5 ACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGG
               || |||||  ||   ||| |  | :  || |   | | |  | | || ||
Db         307 ACACAGCCAGGCTAGATTATCTCGYTGTCATAATCTTTGCAAAGGCAGTTTC

Qy          65 AACAGCCANCA 75
               ||:|||||  ||
Db         247 AAYAGCCACCA 237


RESULT 15
US-10-057-810-41
; Sequence 41, Application US/10057810
; Publication No. US20030143547A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Xianqiang
;  TITLE OF INVENTION: METHOD FOR IDENTIFYING MULTIPLE ACTIVATED T
;  FILE REFERENCE: 26757-709
;  CURRENT APPLICATION NUMBER: US/10/057,810
;  CURRENT FILING DATE:  2002-01-24
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PP11 Reporter Sequence
US-10-057-810-41

  Query Match             36.9%;  Score 28.4;  DB 14;  Length 100;
  Best Local Similarity   96.7%;  Pred. No. 0.37;
```

```
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
  |||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 76



1
Application US/10057828
o. US20030148287A1
MATION:
i, Xiangqiang
Jiang, Xin
ENTION: LIBRARIES AND KITS FOR DETECTING TRANSCRIPTION FACTOR ACTIVITY
CE: 26757-707
ICATION NUMBER: US/10/057,828
NG DATE:  2002-01-24
Q ID NOS: 60
tentIn version 3.1


rtificial Sequence

MATION: PP11 Reporter Sequence
1

          36.9%;  Score 28.4;  DB 14;  Length 100;
milarity   96.7%;  Pred. No. 0.37;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
  |||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 76



4411
1, Application US/10369493
o. US20030233675A1
MATION:
ao, Yongwei
Hinkle, Gregory J.
Slater, Steven C.
Goldman, Barry S.
Chen, Xianfeng
ENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
ENTION:  PLANTS WITH IMPROVED PROPERTIES
CE: 38-10(52052)B
ICATION NUMBER: US/10/369,493
NG DATE:  2003-02-28
ATION NUMBER: US 60/360,039
 DATE: 2002-02-21
Q ID NOS: 47374
11
5

scherichia coli
4411

          36.9%;  Score 28.4;  DB 15;  Length 3075;
milarity   96.7%;  Pred. No. 0.93;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
  |||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1103



pplication US/10410012
; Publication No. US20030228276A1
; GENERAL INFORMATION:
;  APPLICANT: Pitossi, Fernando  J
;  APPLICANT:  Eisel, Ulrich L M
;  TITLE OF INVENTION: Neuroprotective and neurodegenerative effec
;  TITLE OF INVENTION:  in the substantia nigra and a new model fo
;  FILE REFERENCE: 1418P US/101141-17
;  CURRENT APPLICATION NUMBER: US/10/410,012
;  CURRENT FILING DATE:  2003-04-09
;  PRIOR APPLICATION NUMBER: US 60/370,974
;  PRIOR FILING DATE: 2002-04-09
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 3078
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-410-012-8

  Query Match           36.9%;  Score 28.4;  DB 15;  Length 3078;
  Best Local Similarity   96.7%;  Pred. No. 0.93;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | |||||||||||||||||||||||||||||
Db       1071 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1100


RESULT 19
US-10-162-688-1
; Sequence 1, Application US/10162688
; Publication No. US20030100528A1
; GENERAL INFORMATION:
;  APPLICANT: INSTITUT PASTEUR
;  APPLICANT:  CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
;  TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDES HAVING A REDUCED OF
;  TITLE OF INVENTION:  INCREASED CONTENT OF EPIGENETIC CONTROL MOT
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: B4752_IP
;  CURRENT APPLICATION NUMBER: US/10/162,688
;  CURRENT FILING DATE:  2002-06-06
;  PRIOR APPLICATION NUMBER: CANADA NO 2,291,367
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 3150
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(3150)
;   OTHER INFORMATION: Nucleic acid sequence of LacZ
US-10-162-688-1

  Query Match           36.9%;  Score 28.4;  DB 14;  Length 3150;
  Best Local Similarity   96.7%;  Pred. No. 0.93;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | |||||||||||||||||||||||||||||
Db       1149 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1178


RESULT 20
US-09-965-703-74
; Sequence 74, Application US/09965703
; Patent No. US20020119521A1
; GENERAL INFORMATION:
;  APPLICANT: Rohm and Haas Company
;  APPLICANT:  Palli, Subba Reddy
;  APPLICANT:  Kapitskaya, Marianna Zinovjevna
;  APPLICANT:  Cress, Dean Ervin
```

ENTION: No. US20020119521A1el Ecdysone Receptor-Based Inducible Gene Expr
CE: A01020B
ICATION NUMBER: US/09/965,703
NG DATE: 2001-09-26
ATION NUMBER: 60/191,355
DATE: 2000-03-22
ATION NUMBER: 60/269,799
DATE: 2001-02-20
ATION NUMBER: PCT/US01/09050
DATE: 2001-03-21
Q ID NOS: 75
tentIn version 3.1

7

scherichia coli
4

36.9%; Score 28.4; DB 9; Length 3157;
milarity 96.7%; Pred. No. 0.93;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1199

186
Application US/10157305A
No. US20030166099A1
RMATION:
Sabbadini, Roger A.
Surber, Mark W.
Neil Berkley
Anca M. Segall
Robert Klepper
ENTION: MINICELL COMPRISING MEMBRANE PROTEINS
ICE: MPEX.008DV1
LICATION NUMBER: US/10/157,305A
NG DATE: 2002-05-28
CATION NUMBER: 60/293,566
DATE: 2001-05-24
CATION NUMBER: 60/359,843
DATE: 2002-02-25
CATION NUMBER: 10/154,951
DATE: 2002-05-24
EQ ID NOS: 258
astSEQ for Windows Version 4.0
5
34

Artificial Sequence

RMATION: Gene encoding a fusion protein
-186

36.9%; Score 28.4; DB 14; Length 3284;
imilarity 96.7%; Pred. No. 0.94;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

186
, Application US/10157391
No. US20030166279A1
RMATION:
Sabbadini, Roger A.
Neil Berkley

; TITLE OF INVENTION: MINICELL-BASED TRANSFECTION
; FILE REFERENCE: MPEX.008DV14
; CURRENT APPLICATION NUMBER: US/10/157,391
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-10-157-391-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284;
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 23
US-10-157-096-186
; Sequence 186, Application US/10157096
; Publication No. US20030190601A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: TARGET DISPLAY ON MINICELLS
; FILE REFERENCE: MPEX.008DV12
; CURRENT APPLICATION NUMBER: US/10/157,096
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,51
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-10-157-096-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 24
US-10-157-302-186
; Sequence 186, Application US/10157302
; Publication No. US20030190683A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Surber, Mark W.

/ENTION: MINICELL-BASED RATIONAL DRUG DESIGN
ICE: MPEX.008DV17
LICATION NUMBER: US/10/157,302
ING DATE: 2002-10-01
CATION NUMBER: 60/293,566
G DATE: 2001-05-24
CATION NUMBER: 60/359,843
G DATE: 2002-02-25
CATION NUMBER: 10/154,951
G DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0
;
4

rtificial Sequence

MATION: Fusion protein
86

36.9%; Score 28.4; DB 14; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

186
Application US/10157215A
o. US20030190749A1
MATION:
urber, Mark W.
Sabbadini, Roger A.
Segall, Anca M.
Berkley, Neil
ENTION: MINICELL-PRODUCING PARENT CELLS
ICE: MPEX.008DV23
ICATION NUMBER: US/10/157,215A
NG DATE: 2002-05-28
ATION NUMBER: 60/293,566
DATE: 2001-05-24
ATION NUMBER: 60/359,843
DATE: 2002-02-25
ATION NUMBER: 10/154,951
DATE: 2002-05-24
Q ID NOS: 258
stSEQ for Windows Version 4.0

4

rtificial Sequence

MATION: Gene encoding a fusion protein
186

36.9%; Score 28.4; DB 14; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

86
Application US/10157299
o. US20030194714A1
MATION:
abbadini, Roger A.

; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.
; TITLE OF INVENTION: MINICELL-BASED TRANSFORMATION
; FILE REFERENCE: MPEX.008DV15
; CURRENT APPLICATION NUMBER: US/10/157,299
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-10-157-299-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 27
US-10-154-951B-186
; Sequence 186, Application US/10154951B
; Publication No. US20030194798A1
; GENERAL INFORMATION:
; APPLICANT: Surber, Mark W.
; APPLICANT: Sabbadini, Roger A.
; TITLE OF INVENTION: MINICELL COMPOSITIONS AND METHODS
; FILE REFERENCE: MPEX.008A
; CURRENT APPLICATION NUMBER: US/10/154,951B
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gene encoding a fusion protein
US-10-154-951B-186

Query Match 36.9%; Score 28.4; DB 14; Length 3284;
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 28
US-10-156-831-186
; Sequence 186, Application US/10156831
; Publication No. US20030198995A1
; GENERAL INFORMATION:
; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Neil Berkley
; APPLICANT: Surber, Mark W.

```
TENTION: FORWARD SCREENING USING MINICELLS
ICE: MPEX.008DV20
ICATION NUMBER: US/10/156,831
NG DATE:  2002-10-02
ATION NUMBER: 60/293,566
 DATE: 2001-05-24
ATION NUMBER: 60/359,843
 DATE: 2002-02-25
ATION NUMBER: 10/154,951
 DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0
;
4

rtificial Sequence

MATION: Fusion protein
86

          36.9%;  Score 28.4;  DB 14;  Length 3284;
milarity   96.7%;  Pred. No. 0.94;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283


86
  Application US/10157147
o. US20030198996A1
MATION:
urber, Mark W.
 Berkley, Neil
 Gerhart, William
 Sabbadini, Roger A.
ENTION: MINICELL LIBRARIES
ICE: MPEX.008DV19
ICATION NUMBER: US/10/157,147
NG DATE:  2002-05-28
ATION NUMBER: 60/293,566
 DATE: 2001-05-24
ATION NUMBER: 60/359,843
 DATE: 2002-02-25
ATION NUMBER: 10/154,951
 DATE: 2002-05-24
Q ID NOS: 258
stSEQ for Windows Version 4.0
;
4

rtificial Sequence

MATION: Gene encoding a fusion protein
86

          36.9%;  Score 28.4;  DB 14;  Length 3284;
milarity   96.7%;  Pred. No. 0.94;
;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283


86
, Application US/10157166
o. US20030199005A1
MATION:
abbadini, Roger A.
;  APPLICANT:  Robert Klepper
;  TITLE OF INVENTION: SOLID SUPPORTS WITH MINICELLS
;  FILE REFERENCE: MPEX.008DV4
;  CURRENT APPLICATION NUMBER: US/10/157,166
;  CURRENT FILING DATE:  2002-05-28
;  PRIOR APPLICATION NUMBER: 60/293,566
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: 60/359,843
;  PRIOR FILING DATE: 2002-02-25
;  PRIOR APPLICATION NUMBER: 10/154,951
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 257
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
;   LENGTH: 3284
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-157-166-186

  Query Match             36.9%;  Score 28.4;  DB 14;  Length 3284;
  Best Local Similarity   96.7%;  Pred. No. 0.94;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | ||||||||||||||||||||||||||||
Db       1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283


RESULT 31
US-10-156-902-186
; Sequence 186, Application US/10156902
; Publication No. US20030199088A1
; GENERAL INFORMATION:
;  APPLICANT: Sabbadini, Roger A.
;  APPLICANT:  Neil Berkley
;  APPLICANT:  Surber, Mark W.
;  TITLE OF INVENTION: MINICELL-BASED GENE THERAPY
;  FILE REFERENCE: MPEX.008DV16
;  CURRENT APPLICATION NUMBER: US/10/156,902
;  CURRENT FILING DATE:  2002-05-28
;  PRIOR APPLICATION NUMBER: 60/293,566
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: 60/359,843
;  PRIOR FILING DATE: 2002-02-25
;  PRIOR APPLICATION NUMBER: 10/154,951
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 257
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
;   LENGTH: 3284
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion protein
US-10-156-902-186

  Query Match             36.9%;  Score 28.4;  DB 14;  Length 3284;
  Best Local Similarity   96.7%;  Pred. No. 0.94;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | ||||||||||||||||||||||||||||
Db       1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283


RESULT 32
US-10-157-318-186
; Sequence 186, Application US/10157318
; Publication No. US20030199089A1
; GENERAL INFORMATION:
;  APPLICANT: Surber, Mark W.
```

Sabbadini, Roger A.
ENTION: MEMBRANE TO MEMBRANE DELIVERY
CE: MPEX.008DV8
ICATION NUMBER: US/10/157,318
NG DATE: 2002-05-28
ATION NUMBER: 60/293,566
DATE: 2001-05-24
ATION NUMBER: 60/359,843
DATE: 2002-02-25
ATION NUMBER: 10/154,51
DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0

4

rtificial Sequence

MATION: Fusion protein
86

36.9%; Score 28.4; DB 14; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

86
Application US/10156811
o. US20030207833A1
MATION:
eil Berkley
Robert Klepper
Sabbadini, Roger A.
ENTION: PHARMACEUTICAL COMPOSITIONS WITH
ENTION: MINICELLS
CE: MPEX.008DV9
ICATION NUMBER: US/10/156,811
NG DATE: 2002-05-28
ATION NUMBER: 60/293,566
DATE: 2001-05-24
ATION NUMBER: 60/359,843
DATE: 2002-02-25
ATION NUMBER: 10/154,951
DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0

4

rtificial Sequence

MATION: Fusion protein
86

36.9%; Score 28.4; DB 15; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

186
Application US/10157320A
o. US20030219408A1
MATION:

; APPLICANT: Sabbadini, Roger A.
; APPLICANT: Klepper, Robert
; TITLE OF INVENTION: METHODS OF MAKING PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITIONS WITH MINICELLS
; FILE REFERENCE: MPEX.008DV10
; CURRENT APPLICATION NUMBER: US/10/157,320A
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gene encoding a fusion protein
US-10-157-320A-186

Query Match 36.9%; Score 28.4; DB 15; Length 3284;
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 35
US-10-157-418A-186
; Sequence 186, Application US/10157418A
; Publication No. US20030219888A1
; GENERAL INFORMATION:
; APPLICANT: Segall, Anca M.
; APPLICANT: Klepper, Robert
; TITLE OF INVENTION: MINICELL-BASED BIOREMEDIATION
; FILE REFERENCE: MPEX.008DV22
; CURRENT APPLICATION NUMBER: US/10/157,418A
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 10/154,951
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/359,843
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/293,566
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 3284
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gene encoding a fusion protein
US-10-157-418A-186

Query Match 36.9%; Score 28.4; DB 15; Length 3284;
Best Local Similarity 96.7%; Pred. No. 0.94;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 36
US-10-157-317-186
; Sequence 186, Application US/10157317
; Publication No. US20030232335A1
; GENERAL INFORMATION:

Burber, Mark W.
Sabbadini, Roger A.
Neil Berkley
/ENTION: MINICELL-BASED SCREENING FOR COMPOUNDS
/ENTION: AND PROTEINS THAT MODULATE THE ACTIVITY OF SIGNALLING
/ENTION: PROTEINS
NCE: MPEX.008DV21
LICATION NUMBER: US/10/157,317
ING DATE: 2002-05-28
CATION NUMBER: 60/293,566
DATE: 2001-05-24
CATION NUMBER: 60/359,843
DATE: 2002-02-25
CATION NUMBER: 10/154,951
DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0
;
4

rtificial Sequence

MATION: Fusion protein
.86

36.9%; Score 28.4; DB 15; Length 3284;
milarity 96.7%; Pred. No. 0.94;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

.86
Application US/10157339
o. US20040005700A1
MATION:
Burber, Mark W.
Matthew Giacalone
/ENTION: POROPLASTS
NCE: MPEX.008DV3
LICATION NUMBER: US/10/157,339
NG DATE: 2002-08-29
CATION NUMBER: 60/293,566
DATE: 2001-05-24
CATION NUMBER: 60/359,843
DATE: 2002-02-25
CATION NUMBER: 10/154,951
DATE: 2002-05-24
Q ID NOS: 257
stSEQ for Windows Version 4.0
;
4

rtificial Sequence

MATION: Fusion protein
186

36.9%; Score 28.4; DB 15; Length 3284;
milarity 96.7%; Pred. No. 0.94;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

2
Application US/10338411
; Publication No. US20030153045A1
; GENERAL INFORMATION:
; APPLICANT: Butt, Tauseef
; APPLICANT: Weeks, Stephen
; APPLICANT: Tran, Hiep
; APPLICANT: Malakhova, Oxana
; APPLICANT: Malakhov, Micheal
; TITLE OF INVENTION: Methods and Compositions for Protein
; TITLE OF INVENTION: Expression and Purification
; FILE REFERENCE: 1955-2792US1
; CURRENT APPLICATION NUMBER: US/10/338,411
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/346,449
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-338-411-32

Query Match 36.9%; Score 28.4; DB 14; Length 3396;
Best Local Similarity 96.7%; Pred. No. 0.95;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
Db 1392 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1421

RESULT 39
US-10-389-640-32
; Sequence 32, Application US/10389640
; Publication No. US20040018591A1
; GENERAL INFORMATION:
; APPLICANT: Butt, Tauseef
; APPLICANT: Weeks, Stephen
; APPLICANT: Tran, Hiep
; APPLICANT: Malakhova, Oxana
; APPLICANT: Malakhov, Micheal
; TITLE OF INVENTION: Methods and Compositions for Protein
; TITLE OF INVENTION: Expression and Purification
; FILE REFERENCE: 1955-P02972US2
; CURRENT APPLICATION NUMBER: US/10/389,640
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 10/338,411
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/346,449
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-389-640-32

Query Match 36.9%; Score 28.4; DB 15; Length 3396;
Best Local Similarity 96.7%; Pred. No. 0.95;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
Db 1392 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1421

RESULT 40
US-09-956-998A-1

pplication US/09956998A
20020082236A1
MATION:
lack Jr., Charles A.
ENTION: COMPOSITIONS AND METHODS FOR ACTIVATING
ENTION: GENES OF INTEREST
CE: 5722-2(35722/191928)
ICATION NUMBER: US/09/956,998A
NG DATE: 2001-09-20
ATION NUMBER: 09/446,402
DATE: 1999-12-20
ATION NUMBER: 60/050,772
DATE: 1997-06-25
Q ID NOS: 19
stSEQ for Windows Version 4.0

9

rtificial Sequence

MATION: Recombinant Molecule containing multiple cloning
MATION: site, kozak sequence, LacZ gene.
isc_feature
1)...(64)
MATION: Multiple cloning site
isc_feature
65)...(79)
MATION: Consensus sequence for the "Kozak sequence"
MATION: (translation initiation)
rim_transcript
80)...(4279)
MATION: Beta galactosidase
1

36.9%; Score 28.4; DB 9; Length 4279;
milarity 96.7%; Pred. No. 1;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1864

d: April 13, 2004, 21:46:12
5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

icleic search, using sw model

April 13, 2004, 18:19:17 ; Search time 43 Seconds
(without alignments)
993.749 Million cell updates/sec

US-10-090-326-25
77
1 acttacncagcccagcanca.....ggagnacaacagccancagt 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

682709 seqs, 277475446 residues

hits satisfying chosen parameters: 1365418

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

is the number of results predicted by chance to have a
ater than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

	% Query Match	Length	DB	ID	Description
	39.0	1059	2	US-08-484-956-58	Sequence 58, Appl
	39.0	1059	2	US-08-484-956-59	Sequence 59, Appl
	39.0	1059	2	US-08-757-653-58	Sequence 58, Appl
	39.0	1059	2	US-08-757-653-59	Sequence 59, Appl
	39.0	1059	4	US-08-520-946-58	Sequence 58, Appl
	39.0	1059	4	US-08-520-946-59	Sequence 59, Appl
	39.0	1059	4	US-09-655-378A-58	Sequence 58, Appl
	39.0	1059	4	US-09-655-378A-59	Sequence 59, Appl
	39.0	1587	2	US-08-484-956-60	Sequence 60, Appl
	39.0	1587	2	US-08-484-956-61	Sequence 61, Appl
	39.0	1587	2	US-08-757-653-60	Sequence 60, Appl
	39.0	1587	2	US-08-757-653-61	Sequence 61, Appl
	39.0	1587	4	US-08-520-946-60	Sequence 60, Appl
	39.0	1587	4	US-08-520-946-61	Sequence 61, Appl
	39.0	1587	4	US-09-655-378A-60	Sequence 60, Appl
	39.0	1587	4	US-09-655-378A-61	Sequence 61, Appl
	39.0	1894	1	US-08-233-305-1	Sequence 1, Appli
	39.0	1894	1	US-08-203-054-1	Sequence 1, Appli
	39.0	1894	1	US-08-636-676-1	Sequence 1, Appli
	39.0	1894	2	US-08-545-212-1	Sequence 1, Appli
	39.0	1894	2	US-08-583-238-1	Sequence 1, Appli
	39.0	1905	3	US-08-540-922D-1	Sequence 1, Appli
	39.0	1906	1	US-08-081-673-1	Sequence 1, Appli
	39.0	1906	1	US-08-587-391-1	Sequence 1, Appli
	39.0	1906	4	US-10-011-436-1	Sequence 1, Appli
	39.0	1910	2	US-08-370-909-18	Sequence 18, Appl
	39.0	2384	1	US-08-181-471-1	Sequence 1, Appli
	28	30	39.0	2384	3	US-09-056-105-3	Sequenc
	29	28.4	36.9	3096	4	US-09-489-039A-4015	Sequenc
	30	28.4	36.9	3152	1	US-07-924-028A-3	Sequenc
	31	28.4	36.9	3365	1	US-07-789-915A-1	Sequenc
	32	28.4	36.9	3365	1	US-08-005-002C-1	Sequenc
	33	28.4	36.9	3365	1	US-08-487-203A-1	Sequenc
	34	28.4	36.9	4279	4	US-09-446-402A-1	Sequenc
	35	28.4	36.9	4810	3	US-08-852-629-11	Sequenc
	36	28.4	36.9	4838	3	US-08-852-629-15	Sequenc
	37	28.4	36.9	6700	4	US-09-654-449-1	Sequenc
	38	28.4	36.9	7252	4	US-09-238-356-27	Sequenc
	39	28.4	36.9	7366	6	5169760-3	Patent I
c	40	28.4	36.9	7387	4	US-09-238-356-28	Sequenc
	41	28.4	36.9	7408	4	US-09-115-475-13	Sequenc
c	42	28.4	36.9	7612	4	US-09-700-934A-2	Sequenc
	43	28.4	36.9	7664	4	US-08-994-689C-10	Sequenc
	44	28.4	36.9	7852	3	US-08-836-022A-2	Sequenc
	45	28.4	36.9	7852	3	US-09-427-048A-2	Sequenc

ALIGNMENTS

RESULT 1
US-08-484-956-58
; Sequence 58, Application US/08484956
; Patent No. 5843654
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: OLDENBURG, MARY C.
; APPLICANT: HEISLER, LAURA
; TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,956
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,601
; FILING DATE: 09-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,164
; FILING DATE: 09-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/254,359
; FILING DATE: 06-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,384
; FILING DATE: 04-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/986,330
; FILING DATE: 07-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CARRPOLL], PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338

```
FOR SEQ ID NO:  58:
HARACTERISTICS:
 1059 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  DNA (genomic)
8

          39.0%;  Score 30;  DB 2;  Length 1059;
milarity   67.9%;  Pred. No. 0.026;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   || |  |   | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 552



9
Application US/08484956
43654
RMATION:
  DAHLBERG, JAMES E.
  LYAMICHEV, VICTOR I.
  BROW, MARY ANN D.
  OLDENBURG, MARY C.
  HEISLER, LAURA
NVENTION:  DETECTION OF p53 MUTATIONS
SEQUENCES:  114
ENCE ADDRESS:
E:  HAVERSTOCK, MEDLEN & CARROLL
 220 MONTGOMERY STREET, SUITE 2200
AN FRANCISCO
CALIFORNIA
  UNITED STATES OF AMERICA
104
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/484,956
ATE:
CATION:  435
ICATION DATA:
ION NUMBER:  US 08/402,601
ATE:  09-MAR-1995
ICATION DATA:
ION NUMBER:  US 08/337,164
ATE:  09-NOV-1994
ICATION DATA:
ION NUMBER:  US 08/254,359
ATE:  06-JUN-1994
ICATION DATA:
ION NUMBER:  US 08/073,384
ATE:  04-JUN-1993
ICATION DATA:
ION NUMBER:  US 07/986,330
ATE:  07-DEC-1992
AGENT INFORMATION:
ARRPOLL ], PETER G.
ATION NUMBER:  32,837
E/DOCKET NUMBER:  FORS-01801
NICATION INFORMATION:
NE:  (415) 705-8410
:  (415) 397-8338
 FOR SEQ ID NO:  59:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1059 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-484-956-59

  Query Match            39.0%;  Score 30;  DB 2;  Length 1059;
  Best Local Similarity  67.9%;  Pred. No. 0.026;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        535 AGTACAACAGCCATCAGT 552


RESULT 3
US-08-757-653-58
; Sequence 58, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acid Using
;    TITLE OF INVENTION:  Thermostable FEN-1 Endonucleases
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02565
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1059 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-757-653-58

  Query Match            39.0%;  Score 30;  DB 2;  Length 1059;
  Best Local Similarity  67.9%;  Pred. No. 0.026;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
```

```
| ||||||||||| ||||
GTACAACAGCCATCAGT 552



9
Application US/08757653
43669
RMATION:
  Kaiser, Michael W.
  Lyamichev, Victor I.
  Lyamichev, Natasha
NVENTION:  Cleavage Of Nucleic Acid Using
NVENTION:  Thermostable FEN-1 Endonucleases
SEQUENCES:  190
ENCE ADDRESS:
E:  Medlen & Carroll, LLP
 220 Montgomery Street, Suite 2200
an Francisco
California
  United States Of America
104
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.30
PLICATION DATA:
ION NUMBER:  US/08/757,653
ATE:
CATION:  435
GENT INFORMATION:
ngolia, Diane E.
TION NUMBER:  40,027
E/DOCKET NUMBER:  FORS-02565
ICATION INFORMATION:
E:  (415) 705-8410
  (415) 397-8338
FOR SEQ ID NO:  59:
HARACTERISTICS:
 1059 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  DNA (genomic)
9

          39.0%;  Score 30;  DB 2;  Length 1059;
nilarity   67.9%;  Pred. No. 0.026;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 552



3
Application US/08520946
72424
RMATION:
  BROW, MARY ANN D.
  LYAMICHEV, VICTOR I.
  OLIVE, DAVID M.
NVENTION:  RAPID DETECTION AND IDENTIFICATION OF
NVENTION:  PATHOGENS
SEQUENCES:  160
ENCE ADDRESS:
E:  MEDLEN & CARROLL
;       STREET:  220 MONTGOMERY STREET, SUITE 2200
;       CITY:  SAN FRANCISCO
;       STATE:  CALIFORNIA
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/520,946
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CARROLL, PETER G.
;       REGISTRATION NUMBER:  32,837
;       REFERENCE/DOCKET NUMBER:  FORS-01756
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 705-8410
;       TELEFAX:  (415) 397-8338
;   INFORMATION FOR SEQ ID NO:  58:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1059 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-520-946-58

  Query Match             39.0%;  Score 30;  DB 4;  Length 1059;
  Best Local Similarity   67.9%;  Pred. No. 0.026;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
                || |||||||||| ||||
Db        535 AGTACAACAGCCATCAGT 552


RESULT 6
US-08-520-946-59
; Sequence 59, Application US/08520946
; Patent No. 6372424
;  GENERAL INFORMATION:
;     APPLICANT:  BROW, MARY ANN D.
;     APPLICANT:  LYAMICHEV, VICTOR I.
;     APPLICANT:  OLIVE, DAVID M.
;     TITLE OF INVENTION:  RAPID DETECTION AND IDENTIFICATION OF
;     TITLE OF INVENTION:  PATHOGENS
;     NUMBER OF SEQUENCES:  160
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MEDLEN & CARROLL
;       STREET:  220 MONTGOMERY STREET, SUITE 2200
;       CITY:  SAN FRANCISCO
;       STATE:  CALIFORNIA
;       COUNTRY:  UNITED STATES OF AMERICA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/520,946
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CARROLL, PETER G.
;       REGISTRATION NUMBER:  32,837
```

:E/DOCKET NUMBER: FORS-01756
:ICATION INFORMATION:
:E: (415) 705-8410
(415) 397-8338
FOR SEQ ID NO: 59:
:HARACTERISTICS:
1059 base pairs
:ucleic acid
)NESS: single
:: linear
:YPE: DNA (genomic)
:9

39.0%; Score 30; DB 4; Length 1059;
.milarity 67.9%; Pred. No. 0.026;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

:CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
:CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

:GNACAACAGCCANCAGT 77
:GTACAACAGCCATCAGT 552

:58
Application US/09655378A
:73616
:ORMATION:
:ANT: BROW, MARY ANN D.
LYAMICHEV, VICTOR I.
OLIVE, DAVID M.
OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
PATHOGENS
: OF SEQUENCES: 165
:PONDENCE ADDRESS:
:DDRESSEE: MEDLEN & CARROLL
:TREET: 220 MONTGOMERY STREET, SUITE 2200
:ITY: SAN FRANCISCO
:TATE: CALIFORNIA
:OUNTRY: UNITED STATES OF AMERICA
:IP: 94104
:ER READABLE FORM:
:EDIUM TYPE: Floppy disk
:OMPUTER: IBM PC compatible
)PERATING SYSTEM: PC-DOS/MS-DOS
:OFTWARE: PatentIn Release #1.0, Version #1.30
:T APPLICATION DATA:
:PPLICATION NUMBER: US/09/655,378A
:ILING DATE: 05-Sep-2000
:LASSIFICATION: <Unknown>
:EY/AGENT INFORMATION:
:AME: CARROLL, PETER G.
:EGISTRATION NUMBER: 32,837
:EFERENCE/DOCKET NUMBER: FORS-01756
)MMUNICATION INFORMATION:
:ELEPHONE: (415) 705-8410
:ELEFAX: (415) 397-8338
: FOR SEQ ID NO: 58:
:CE CHARACTERISTICS:
:ENGTH: 1059 base pairs
:YPE: nucleic acid
:TRANDEDNESS: double
:OPOLOGY: linear
:LE TYPE: DNA (genomic)
:CE DESCRIPTION: SEQ ID NO: 58:
-58

39.0%; Score 30; DB 4; Length 1059;
imilarity 67.9%; Pred. No. 0.026;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
Db 535 AGTACAACAGCCATCAGT 552

RESULT 8
US-09-655-378A-59
; Sequence 59, Application US/09655378A
; Patent No. 6673616
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/655,378A
; FILING DATE: 05-Sep-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-655-378A-59

Query Match 39.0%; Score 30; DB 4; Length 1059;
Best Local Similarity 67.9%; Pred. No. 0.026;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
Db 535 AGTACAACAGCCATCAGT 552

RESULT 9
US-08-484-956-60
; Sequence 60, Application US/08484956
; Patent No. 5843654
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.

```
   LYAMICHEV, VICTOR I.
   BROW, MARY ANN D.
   OLDENBURG, MARY C.
   HEISLER, LAURA
NVENTION:  DETECTION OF p53 MUTATIONS
SEQUENCES:  114
)ENCE ADDRESS:
:E:  HAVERSTOCK, MEDLEN & CARROLL
  220 MONTGOMERY STREET, SUITE 2200
:AN FRANCISCO
 CALIFORNIA
   UNITED STATES OF AMERICA
104
:EADABLE FORM:
'YPE:  Floppy disk
::  IBM PC compatible
'G SYSTEM:  PC-DOS/MS-DOS
::  PatentIn Release #1.0, Version #1.25
'PLICATION DATA:
'ION NUMBER:  US/08/484,956
'ATE:
CATION:  435
.ICATION DATA:
'ION NUMBER:  US 08/402,601
'ATE:  09-MAR-1995
.ICATION DATA:
'ION NUMBER:  US 08/337,164
'ATE:  09-NOV-1994
.ICATION DATA:
'ION NUMBER:  US 08/254,359
'ATE:  06-JUN-1994
.ICATION DATA:
'ION NUMBER:  US 08/073,384
ATE:  04-JUN-1993
.ICATION DATA:
ION NUMBER:  US 07/986,330
ATE:  07-DEC-1992
GENT INFORMATION:
ARRPOLL ], PETER G.
TION NUMBER:  32,837
E/DOCKET NUMBER:  FORS-01801
ICATION INFORMATION:
E:  (415) 705-8410
   (415) 397-8338
FOR SEQ ID NO:  60:
HARACTERISTICS:
 1587 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  DNA (genomic)
0

          39.0%;  Score 30;  DB 2;  Length 1587;
milarity   67.9%;  Pred. No. 0.03;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  || |   | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 859



1
Application US/08484956
43654
RMATION:
  DAHLBERG, JAMES E.
  LYAMICHEV, VICTOR I.
```

```
;    APPLICANT:  BROW, MARY ANN D.
;    APPLICANT:  OLDENBURG, MARY C.
;    APPLICANT:  HEISLER, LAURA
;    TITLE OF INVENTION:  DETECTION OF p53 MUTATIONS
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAVERSTOCK, MEDLEN & CARROLL
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,956
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/402,601
;      FILING DATE:  09-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/337,164
;      FILING DATE:  09-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/254,359
;      FILING DATE:  06-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/073,384
;      FILING DATE:  04-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/986,330
;      FILING DATE:  07-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARRPOLL ], PETER G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  FORS-01801
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1587 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-484-956-61

  Query Match           39.0%;  Score 30;  DB 2;  Length 1587;
  Best Local Similarity   67.9%;  Pred. No. 0.03;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  || |   | | |
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 11
US-08-757-653-60
; Sequence 60, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
```

INVENTION: Cleavage Of Nucleic Acid Using
INVENTION: Thermostable FEN-1 Endonucleases
SEQUENCES: 190
DENCE ADDRESS:
EE: Medlen & Carroll, LLP
220 Montgomery Street, Suite 2200
San Francisco
California
: United States Of America
4104
READABLE FORM:
TYPE: Floppy disk
R: IBM PC compatible
NG SYSTEM: PC-DOS/MS-DOS
E: PatentIn Release #1.0, Version #1.30
PPLICATION DATA:
TION NUMBER: US/08/757,653
DATE:
ICATION: 435
AGENT INFORMATION:
Ingolia, Diane E.
ATION NUMBER: 40,027
CE/DOCKET NUMBER: FORS-02565
NICATION INFORMATION:
NE: (415) 705-8410
: (415) 397-8338
FOR SEQ ID NO: 60:
CHARACTERISTICS:
1587 base pairs
nucleic acid
DNESS: single
Y: linear
TYPE: DNA (genomic)
50

39.0%; Score 30; DB 2; Length 1587;
imilarity 67.9%; Pred. No. 0.03;
: Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||| | | || | | | ||||||| |
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 859

51
Application US/08757653
343669
ORMATION:
: Kaiser, Michael W.
: Lyamichev, Victor I.
: Lyamichev, Natasha
INVENTION: Cleavage Of Nucleic Acid Using
INVENTION: Thermostable FEN-1 Endonucleases
SEQUENCES: 190
DENCE ADDRESS:
EE: Medlen & Carroll, LLP
220 Montgomery Street, Suite 2200
San Francisco
California
: United States Of America
4104
READABLE FORM:
TYPE: Floppy disk
R: IBM PC compatible
NG SYSTEM: PC-DOS/MS-DOS
E: PatentIn Release #1.0, Version #1.30
PPLICATION DATA:
TION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-757-653-61

Query Match 39.0%; Score 30; DB 2; Length 1587;
Best Local Similarity 67.9%; Pred. No. 0.03;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| ||||||| | | || | | |
Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 842 AGTACAACAGCCATCAGT 859

RESULT 13
US-08-520-946-60
; Sequence 60, Application US/08520946
; Patent No. 6372424
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,946
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1587 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-520-946-60

```
          39.0%;  Score 30;  DB 4;  Length 1587;
milarity   67.9%;  Pred. No. 0.03;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 859



1
Application US/08520946
72424
RMATION:
  BROW, MARY ANN D.
  LYAMICHEV, VICTOR I.
  OLIVE, DAVID M.
NVENTION:  RAPID DETECTION AND IDENTIFICATION OF
NVENTION:  PATHOGENS
SEQUENCES:  160
ENCE ADDRESS:
E:  MEDLEN & CARROLL
 220 MONTGOMERY STREET, SUITE 2200
AN FRANCISCO
CALIFORNIA
  UNITED STATES OF AMERICA
104
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/520,946
ATE:
CATION:  435
ENT INFORMATION:
ARROLL, PETER G.
TION NUMBER:  32,837
E/DOCKET NUMBER:  FORS-01756
ICATION INFORMATION:
E:  (415) 705-8410
  (415) 397-8338
FOR SEQ ID NO:  61:
HARACTERISTICS:
 1587 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  DNA (genomic)
-

          39.0%;  Score 30;  DB 4;  Length 1587;
ilarity   67.9%;  Pred. No. 0.03;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 859



50
Application US/09655378A
```

```
; Patent No. 6673616
;   GENERAL INFORMATION:
;        APPLICANT: BROW, MARY ANN D.
;                   LYAMICHEV, VICTOR I.
;                   OLIVE, DAVID M.
;        TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
;                            PATHOGENS
;        NUMBER OF SEQUENCES: 165
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MEDLEN & CARROLL
;             STREET: 220 MONTGOMERY STREET, SUITE 2200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: UNITED STATES OF AMERICA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/655,378A
;             FILING DATE: 05-Sep-2000
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: CARROLL, PETER G.
;             REGISTRATION NUMBER: 32,837
;             REFERENCE/DOCKET NUMBER: FORS-01756
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 60:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1587 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-655-378A-60

  Query Match            39.0%;  Score 30;  DB 4;  Length 1587;
  Best Local Similarity  67.9%;  Pred. No. 0.03;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 16
US-09-655-378A-61
; Sequence 61, Application US/09655378A
; Patent No. 6673616
;   GENERAL INFORMATION:
;        APPLICANT: BROW, MARY ANN D.
;                   LYAMICHEV, VICTOR I.
;                   OLIVE, DAVID M.
;        TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
;                            PATHOGENS
;        NUMBER OF SEQUENCES: 165
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MEDLEN & CARROLL
;             STREET: 220 MONTGOMERY STREET, SUITE 2200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: UNITED STATES OF AMERICA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
```

```
IEDIUM TYPE: Floppy disk
OMPUTER: IBM PC compatible
)PERATING SYSTEM: PC-DOS/MS-DOS
)OFTWARE: PatentIn Release #1.0, Version #1.30
NT APPLICATION DATA:
APPLICATION NUMBER: US/09/655,378A
'ILING DATE: 05-Sep-2000
LASSIFICATION: <Unknown>
EY/AGENT INFORMATION:
AME: CARROLL, PETER G.
EGISTRATION NUMBER: 32,837
EFERENCE/DOCKET NUMBER: FORS-01756
)MMUNICATION INFORMATION:
ELEPHONE: (415) 705-8410
ELEFAX: (415) 397-8338
 FOR SEQ ID NO: 61:
ICE CHARACTERISTICS:
ENGTH: 1587 base pairs
YPE: nucleic acid
TRANDEDNESS: double
OPOLOGY: linear
LE TYPE: DNA (genomic)
ICE DESCRIPTION: SEQ ID NO: 61:
61

          39.0%;  Score 30;  DB 4;  Length 1587;
milarity   67.9%;  Pred. No. 0.03;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |  | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 859



pplication US/08233305
19117
)RMATION:
  W lfel, Thomas; Van Pel, Aline; Brichard, Vincent;
  Boon-Falleur, Thierry
NVENTION:  ISOLATED, TYROSINASE DERIVED PEPTIDES AND
NVENTION:  USES THEREOF
SEQUENCES:  10
ENCE ADDRESS:
E:  Felfe & Lynch
 805 Third Avenue
ew York City
New York
  USA
022
EADABLE FORM:
YPE:  Diskette, 5.25 inch, 360 kb storage
:  IBM PS/2
G SYSTEM:  PC-DOS
:  Wordperfect
PLICATION DATA:
ION NUMBER:  US/08/233,305
ATE:  26-APRIL-1994
CATION:  514
ICATION DATA:
ION NUMBER:  08/203,054
ATE:  28-FEB-1994
ION NUMBER:  08/081,673
ATE:  23-JUNE-1993
ICATION DATA:
ION NUMBER:  08/054,714
ATE:  28-APRIL-1993
ICATION DATA:
```

```
;      APPLICATION NUMBER:  07/994,928
;      FILING DATE:  22-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5519117man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5360.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-233-305-1

  Query Match             39.0%;  Score 30;  DB 1;  Length 1894;
  Best Local Similarity   67.9%;  Pred. No. 0.032;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| |||||||       |  | |  ||  |   | |
Db        788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
                || |||||||||| ||||
Db        848 AGTACAACAGCCATCAGT 865


RESULT 18
US-08-203-054-1
; Sequence 1, Application US/08203054
; Patent No. 5530096
;  GENERAL INFORMATION:
;    APPLICANT:  W lfel, Thomas; Van Pel, Aline; Brichard, Vincent;
;    APPLICANT:  Boon-Falleur, Thierry
;    TITLE OF INVENTION:  ISOLATED, TYROSINASE DERIVED PEPTIDES ANI
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/203,054
;      FILING DATE:  28-FEB-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/081,673
;      FILING DATE:  23-JUNE-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/054,714
;      FILING DATE:  28-APRIL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/994,928
;      FILING DATE:  22-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5530096man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 360
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
```

HARACTERISTICS:
1894 base pairs
ucleic acid
NESS: single
: linear

39.0%; Score 30; DB 1; Length 1894;
milarity 67.9%; Pred. No. 0.032;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 865
```

pplication US/08636676
47271
RMATION:
Boon-Falleur, Thierry; Brichard, Vincent; Van Pel,
Aline; DePlaen, Etienne; Coulie, Pierre; Renauld, Jean Christophe;
Wolfel, Thomas
NVENTION: METHOD FOR IDENTIFYING INDIVIDUALS SUFFERING
NVENTION: FROM A CELLULAR ABNORMALITY SOME OF WHOSE ABNORMAL CELLS PRESE
NVENTION: COMPLEXES OF HLA-A2/TYROSINASE DERIVED PEPTIDES, AND METHODS F
NVENTION: TREATING SAID INDIVIDUALS
SEQUENCES: 1
ENCE ADDRESS:
E: Felfe & Lynch
805 Third Avenue
ew York City
New York
USA
022
EADABLE FORM:
YPE: Diskette, 5.25 inch, 360 kb storage
: IBM PS/2
3 SYSTEM: PC-DOS
: Wordperfect
PLICATION DATA:
ION NUMBER: US/08/636,676
ATE: 23-APR-1996
CATION: 435
ICATION DATA:
ION NUMBER: US/08/295,227
ATE:
ION NUMBER: 07/994,928
ATE: 22-DEC-1992
GENT INFORMATION:
anson, No. 5747271man D.
TION NUMBER: 30,946
E/DOCKET NUMBER: LUD 299
ICATION INFORMATION:
E: (212) 688-9200
(212) 838-3884
FOR SEQ ID NO: 1:
HARACTERISTICS:
1894 base pairs
ucleic acid
NESS: single
: linear

39.0%; Score 30; DB 1; Length 1894;
nilarity 67.9%; Pred. No. 0.032;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
                ||  |||||||||| ||||
Db        848 AGTACAACAGCCATCAGT 865
```

RESULT 20
US-08-545-212-1
; Sequence 1, Application US/08545212
; Patent No. 5843687
; GENERAL INFORMATION:
; APPLICANT: W lfel, Thomas; Van Pel, Aline; Brichard, Vincent;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: ISOLATED, TYROSINASE DERIVED PEPTIDES ANI
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,212
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/233,305
; FILING DATE: 26-APRIL-1994
; CLASSIFICATION: 435
; APPLICATION NUMBER: 08/203,054
; FILING DATE: 28-FEB-1994
; APPLICATION NUMBER: 08/081,673
; FILING DATE: 23-JUNE-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,714
; FILING DATE: 28-APRIL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/994,928
; FILING DATE: 22-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5843687man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5360.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-545-212-1

Query Match 39.0%; Score 30; DB 2; Length 1894;
Best Local Similarity 67.9%; Pred. No. 0.032;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
```

```
|| |||||||||| ||||
AGTACAACAGCCATCAGT 865



1
Application US/08583238
843688
ORMATION:
:  W lfel, Thomas; Van Pel, Aline; Brichard, Vincent;
:  Boon-Falleur, Thierry
INVENTION:  ISOLATED, TYROSINASE DERIVED PEPTIDES AND
INVENTION:  USES THEREOF
 SEQUENCES:  5
DENCE ADDRESS:
EE:  Felfe & Lynch
  805 Third Avenue
New York City
 New York
:  USA
)022
READABLE FORM:
TYPE:  Diskette, 5.25 inch, 360 kb storage
R:  IBM PS/2
NG SYSTEM:  PC-DOS
E:  Wordperfect
PLICATION DATA:
TION NUMBER:  US/08/583,238
DATE:  05-JAN-1996
ICATION:  435
LICATION DATA:
TION NUMBER:  08/203,054
DATE:  28-FEB-1994
ICATION:  435
TION NUMBER:  08/081,673
DATE:  23-JUNE-1993
LICATION DATA:
TION NUMBER:  08/054,714
DATE:  28-APRIL-1993
LICATION DATA:
TION NUMBER:  07/994,928
DATE:  22-DEC-1992
AGENT INFORMATION:
Hanson, No. 5843688man D.
ATION NUMBER:  30,946
E/DOCKET NUMBER:  LUD 360
IICATION INFORMATION:
E:  (212) 688-9200
   (212) 838-3884
 FOR SEQ ID NO:  1:
HARACTERISTICS:
 1894 base pairs
nucleic acid
DNESS:  single
:  linear


          39.0%;  Score 30;  DB 2;  Length 1894;
milarity   67.9%;  Pred. No. 0.032;
   Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |  | |  ||  |  | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 865



-1
Application US/08540922D
```

```
; Patent No. 6284476
;  GENERAL INFORMATION:
;    APPLICANT:  Boon-Falleur, Thierry; Brichard, Vincent; Van
;    APPLICANT:  Pel, Aline; De Plaen, Etienne; Coulie, Pierre;
;    APPLICANT:  Renauld Jean-Christope; Wolfel, Thomas; and
;    APPLICANT:  Lethe, Bernard.
;    TITLE OF INVENTION:  METHOD OF IDENTIFYING INDIVIDUALS SUFFER
;    TITLE OF INVENTION:  FROM A CELLULAR ABNORMALITY SOME OF WHOS
;    TITLE OF INVENTION:  ABNORMAL CELLS PRESENT COMPLEXES OF HUMA
;    TITLE OF INVENTION:  LEUKOCYTE ANTIGEN TYROSINASE DERIVED
;    TITLE OF INVENTION:  PEPTIDES, AND METHODS FOR TREATING SAID
;    TITLE OF INVENTION:  INDIVIDUALS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch 1.44 Mb storage diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/540,922D
;      FILING DATE:  October 11, 1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/054,714
;      FILING DATE:  28 April 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/994,928
;      FILING DATE:  22 December 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5299.5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1905 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-540-922D-1

  Query Match             39.0%;  Score 30;  DB 3;  Length 1905;
  Best Local Similarity   67.9%;  Pred. No. 0.032;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |  | |  ||  |  | |
Db        806 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        866 AGTACAACAGCCATCAGT 883


RESULT 23
US-08-081-673-1
; Sequence 1, Application US/08081673
; Patent No. 5487974
;  GENERAL INFORMATION:
;    APPLICANT:  Boon-Falleur, Thierry; Brichard, Vincent; Van Pel,
;    APPLICANT:  Aline; DePlaen, Etienne; Coulie, Pierre; Renauld,
;    APPLICANT:  Wolfel, Thomas; Lethe, Bernard
;    TITLE OF INVENTION:  METHOD FOR IDENTIFYING INDIVIDUALS SUFFER
;    TITLE OF INVENTION:  FROM A CELLULAR ABNORMALITY SOME OF WHOSE
```

```
NVENTION:  COMPLEXES OF HUMAN LEUKOCYTE ANTIGEN TYROSINASE DERIVED PEPTID
NVENTION:  METHODS FOR TREATING SAID INDIVIDUALS
SEQUENCES:  3
ENCE ADDRESS:
E:  Felfe & Lynch
 805 Third Avenue
ew York City
New York
  USA
022
EADABLE FORM:
YPE:  Diskette, 5.25 inch, 360 kb storage
:  IBM PS/2
G SYSTEM:  PC-DOS
:  Wordperfect
PLICATION DATA:
ION NUMBER:  US/08/081,673
ATE:  23-JUNE-1993
CATION:  435
ICATION DATA:
ION NUMBER:  08/054,714
ATE:  28-APRIL-1992
ION NUMBER:  994,928
ATE:  22-DEC-1992
GENT INFORMATION:
anson, No. 5487974man D.
TION NUMBER:  30,946
E/DOCKET NUMBER:  LUD 299.2
ICATION INFORMATION:
E:  (212) 688-9200
  (212) 838-3884
FOR SEQ ID NO:  1:
HARACTERISTICS:
 1906 base pairs
ucleic acid
NESS:  single
:  linear


          39.0%;  Score 30;  DB 1;  Length 1906;
nilarity   67.9%;  Pred. No. 0.032;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 865

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 883



pplication US/08587391
14316
RMATION:
  Lethe, Bernard; Brichard, Vincent; Van Pel, Aline;
  DePlaen, Etienne; Coulie, Pierre; Renauld, Jean Christophe;
  Wolfel, Thomas; Boon-Falleur, Thierry
NVENTION:  Isolated, Tyrosinase Derived Peptides And
NVENTION:  Uses Thereof.
SEQUENCES:  12
ENCE ADDRESS:
E:  Felfe & Lynch
 805 Third Avenue
ew York City
New York
  USA
022
EADABLE FORM:
YPE:  Diskette, 3.5 inch, 360 kb storage
:  IBM PS/2
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/587,391
;       FILING DATE:  17-January-1996
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/233,305
;       FILING DATE:  26-April-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/203,054
;       FILING DATE:  28-February-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/081,673
;       FILING DATE:  23-June-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/054,714
;       FILING DATE:  28-April-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/994,928
;       FILING DATE:  22-December-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  HANSON, NORMAN
;       REGISTRATION NUMBER:  30, 946
;       REFERENCE/DOCKET NUMBER:  LUD 5431
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-688-9200
;       TELEFAX:  212-838-3884
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1906 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-587-391-1

  Query Match            39.0%;  Score 30;  DB 1;  Length 1906;
  Best Local Similarity  67.9%;  Pred. No. 0.032;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        806 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        866 AGTACAACAGCCATCAGT 883


RESULT 25
US-10-011-436-1
; Sequence 1, Application US/10011436
; Patent No. 6669946
;   GENERAL INFORMATION:
;       APPLICANT: Boon-Falleur, Thierry; Brichard, Vincent; Van P
;                  Aline;
;                  DePlaen, Etienne; Coulie, Pierre; Renauld, Jean
;                  Lethe, Bernard
;       TITLE OF INVENTION: METHOD FOR IDENTIFYING INDIVIDUALS SUF
;                           A
;                           CELLULAR ABNORMALITY SOME OF WHOSE ABN
;                           LEUKOCYTE
;       NUMBER OF SEQUENCES: 13
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fulbright & Jaworski L.L.P.
;            STREET: 801 Pennsylvania Avenue N.W.
;            CITY: Washington
;            STATE: District of Colombia
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.25inch, 1.44MB storage
;            COMPUTER: IBM PS/2
```

)PERATING SYSTEM: PC-DOS
3OFTWARE: Wordperfect
NT APPLICATION DATA:
APPLICATION NUMBER: US/10/011,436
FILING DATE: 11-Dec-2001
CLASSIFICATION: 435
APPLICATION DATA:
APPLICATION NUMBER: 08/511,011
FILING DATE: 5-AUGUST-1995
APPLICATION NUMBER: 08/054,714
FILING DATE: 28-APRIL-1993
APPLICATION NUMBER: 07/994,928
FILING DATE: 22-DECEMBER-1992
NEY/AGENT INFORMATION:
NAME: Mary Anne Schofield
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5299.9 DIV CON
)MMUNICATION INFORMATION:
TELEPHONE: (202) 662-0200
TELEFAX: (202) 662-4643
I FOR SEQ ID NO: 1:
ICE CHARACTERISTICS:
LENGTH: 1906 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ICE DESCRIPTION: SEQ ID NO: 1:
L

39.0%; Score 30; DB 4; Length 1906;
imilarity 67.9%; Pred. No. 0.032;
: Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
'|||||| |||||||||| |||||| | | || | | | ||||||| |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 865

AGNACAACAGCCANCAGT 77
| |||||||||| ||||
AGTACAACAGCCATCAGT 883

.8
Application US/08370909
343648
)RMATION:
ROBBINS, PAUL F.; ROSENBERG,
STEVEN A.
INVENTION: p15 AND TYROSINASE
INVENTION: MELANOMA ANTIGENS AND THEIR USE IN DIAGNOSTIC
INVENTION: AND THERAPEUTIC METHODS
SEQUENCES: 23
)ENCE ADDRESS:
E: MORGAN & FINNEGAN, L.L.P.
345 PARK AVENUE
IEW YORK
NEW YORK
USA
)154
EADABLE FORM:
YPE: FLOPPY DISK
: IBM PC COMPATIBLE
IG SYSTEM: PC-DOS/MS-DOS
: ASCII
)PLICATION DATA:
ION NUMBER: US/08/370,909
)ATE: 10-JAN-1995
CATION: 514
AGENT INFORMATION:
AROL M. GRUPPI
ATION NUMBER: 37,341
E/DOCKET NUMBER: 2026-4155

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1910
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PEPTIDE
US-08-370-909-18

Query Match 39.0%; Score 30; DB 2; Length 1910;
Best Local Similarity 67.9%; Pred. No. 0.032;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | || | | | ||
Db 810 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 870 AGTACAACAGCCATCAGT 887

RESULT 27
US-08-181-471-1
; Sequence 1, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS

: 503..2092

39.0%; Score 30; DB 1; Length 2384;
milarity 67.9%; Pred. No. 0.035;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 1343

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 1361

pplication US/09056105
87569
MATION:
IPPS, THOMAS J.
J, YUNQI
ENTION: VACCINES WITH ENHANCED INTRACELLULAR
ENTION: PROCESSING
CE: 233/221
ICATION NUMBER: US/09/056,105
NG DATE: 1998-04-06
ICATION NUMBER: 60/043,467
NG DATE: 1997-04-10
Q ID NOS: 35
stSEQ for Windows Version 3.0

1

omo sapiens

39.0%; Score 30; DB 3; Length 2384;
ilarity 67.9%; Pred. No. 0.035;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 1343

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 1361

:015
Application US/09489039A
.0836
IATION:
Gary Breton et. al
NTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
NTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
E: 2709.2004001
CATION NUMBER: US/09/489,039A
G DATE: 2000-01-27
ATION NUMBER: US 60/117,747
DATE: 1999-01-29
Q ID NOS: 14342
;
;

ebsiella pneumoniae
:015

36.9%; Score 28.4; DB 4; Length 3096;
ilarity 96.7%; Pred. No. 0.16;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
Db 1095 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1124

RESULT 30
US-07-924-028A-3
; Sequence 3, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCE
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-924-028A-3

Query Match 36.9%; Score 28.4; DB 1; Length 3152;
Best Local Similarity 96.7%; Pred. No. 0.17;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
Db 1086 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1115

RESULT 31
US-07-789-915A-1
; Sequence 1, Application US/07789915A
; Patent No. 5212058
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

```
3E:  Hamilton, Brook, Smith & Reynolds, P.C.
  Two Militia Drive
Lexington
 Massachusetts
:  U.S.A.
2173
READABLE FORM:
YPE:  Floppy disk
R:  IBM PC compatible
NG SYSTEM:  PC-DOS/MS-DOS
E:  PatentIn Release #1.0, Version #1.25
PPLICATION DATA:
TION NUMBER:  US/07/789,915A
DATE:  19911108
CATION:  435
AGENT INFORMATION:
Brook, David E.
ATION NUMBER:  22,592
CE/DOCKET NUMBER:  MIT-5091AA
IICATION INFORMATION:
NE:  617-861-6240
:  617-861-9540
 FOR SEQ ID NO:  1:
CHARACTERISTICS:
  3365 base pairs
NUCLEIC ACID
DNESS:  double
Y:  linear

Y:  CDS
N:  1..3363
:1

           36.9%;  Score 28.4;  DB 1;  Length 3365;
milarity   96.7%;  Pred. No. 0.17;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
  ||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1394



1
pplication US/08005002C
94818
RMATION:
  Baker, Rohan T.
  Tobias, John W.
  Varshavsky, Alexander
NVENTION:  Ubiquitin-Specific Proteases
SEQUENCES:  9
ENCE ADDRESS:
E:  Kevin M. Farrell, P.C.
  P.O. Box 999
ork Harbor
Maine
  U.S.A.
911
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/005,002C
ATE:  15-JAN-1993
CATION:  435
ICATION DATA:
ION NUMBER:  US 07/789,915
ATE:  08-NOV-1991
GENT INFORMATION:
arrell, Kevin M.
;       REGISTRATION NUMBER:  35,505
;       REFERENCE/DOCKET NUMBER:  MIT-5091AAZ
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  207-363-0558
;       TELEFAX:  207-363-0528
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3365 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..3363
US-08-005-002C-1

  Query Match             36.9%;  Score 28.4;  DB 1;  Length 3365;
  Best Local Similarity   96.7%;  Pred. No. 0.17;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | ||||||||||||||||||||||||||||
Db       1365 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1394


RESULT 33
US-08-487-203A-1
; Sequence 1, Application US/08487203A
; Patent No. 5683904
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Rohan T.
;    APPLICANT:  Tobias, John W.
;    APPLICANT:  Varshavsky, Alexander
;    TITLE OF INVENTION:  Ubiquitin-Specific Proteases
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell, P.C.
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  Maine
;      COUNTRY:  U.S.A.
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,203A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/005,002
;      FILING DATE:  15-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Farrell, Kevin M.
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  MIT-5091A3Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  207-363-0558
;      TELEFAX:  207-363-0528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3363
US-08-487-203A-1

  Query Match             36.9%;  Score 28.4;  DB 1;  Length 3365;
```

```
milarity   96.7%;  Pred. No. 0.17;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
 ||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1394



1
pplication US/09446402A
23003
MATION:
lack Jr., Charles A.
ENTION: COMPOSITIONS AND METHODS FOR ACTIVATING
ENTION:  GENES OF INTEREST
CE: 5722-2(35722/191928)
ICATION NUMBER: US/09/446,402A
NG DATE:  1999-12-20
ATION NUMBER: PCT/US98/13093
 DATE: 1998-06-24
ATION NUMBER: 60/050,772
 DATE: 1997-06-25
Q ID NOS: 19
stSEQ for Windows Version 4.0

9

rtificial Sequence

MATION: Recombinant Molecule containing multiple cloning
MATION: site, kozak sequence, LacZ gene.
isc_feature
1)...(64)
MATION: Multiple cloning site
isc_feature
65)...(79)
MATION: Consensus sequence for the "Kozak sequence"
MATION: (translation initiation)
rim_transcript
80)...(4279)
MATION: Beta galactosidase
1

          36.9%;  Score 28.4;  DB 4;  Length 4279;
milarity   96.7%;  Pred. No. 0.19;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
 ||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1864



1
Application US/08852629
06825
RMATION:
  Moyer, Richard W
  Li, Yi
  Hall, Richard L
NVENTION:  ENTOMOPOXVIRUS-VERTEBRATE GENE DELIVERY
NVENTION:  VECTOR AND METHOD
SEQUENCES:  17
ENCE ADDRESS:
E:  Saliwanchik, Lloyd & Saliwanchik
 2421 N.W. 41st Street, Suite A-1
ainesville
Florida
  U.S.A.
506
EADABLE FORM:
YPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/852,629
;       FILING DATE:
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bencen, Gerard H
;       REGISTRATION NUMBER:  35,746
;       REFERENCE/DOCKET NUMBER:  UF-184
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  407-426-7500
;       TELEFAX:  407-839-8589
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4810 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-852-629-11

  Query Match             36.9%;  Score 28.4;  DB 3;  Length 4810;
  Best Local Similarity   96.7%;  Pred. No. 0.19;
  Matches   29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              |||||||||||||||||||||||||||||
Db       2086 CGAGCATCATCCTCTGCATGGTCAGGTCAT 2115


RESULT 36
US-08-852-629-15
; Sequence 15, Application US/08852629
; Patent No. 6106825
;  GENERAL INFORMATION:
;    APPLICANT:  Moyer, Richard W
;    APPLICANT:  Li, Yi
;    APPLICANT:  Hall, Richard L
;    TITLE OF INVENTION:  ENTOMOPOXVIRUS-VERTEBRATE GENE DELIVERY
;    TITLE OF INVENTION:  VECTOR AND METHOD
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  U.S.A.
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,629
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bencen, Gerard H
;      REGISTRATION NUMBER:  35,746
;      REFERENCE/DOCKET NUMBER:  UF-184
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  407-426-7500
;      TELEFAX:  407-839-8589
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4838 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
```

```
:  both
YPE:  cDNA
AL:  NO
:  NO
.5

          36.9%;  Score 28.4;  DB 3;  Length 4838;
milarity   96.7%;  Pred. No. 0.19;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
 |||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 2143



pplication US/09654449
64061
MATION:
almer, Michelle A.J.
Gee, Melissa
Tillotson, Bonnie
Chang, Xiao-Jia
ENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orpha
ENTION:  Receptors by Reporter Enzyme Mutant Complementation
CE: 4085-226-27
ICATION NUMBER: US/09/654,449
NG DATE:  2000-09-01
ATION NUMBER: US 60/180,669
 DATE: 2000-02-07
Q ID NOS: 5
stSEQ for Windows Version 4.0

0

rtificial Sequence

MATION: Nucleotide sequence for pICAST ALC
DS
1457)...(4486)


          36.9%;  Score 28.4;  DB 4;  Length 6700;
milarity   96.7%;  Pred. No. 0.22;
  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
 |||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 2466



7
Application US/09238356
12683
MATION:
Kingsman, et al
ENTION:  Retroviral Vectors
CE:  674523-2006
ICATION NUMBER: US/09/238,356
NG DATE:  1999-01-27
ATION NUMBER:  PCT/GB/03876
 DATE:  1998-12-22
Q ID NOS:  64
PatentIn version 3.0

2

rtificial Sequence, plasmid

isc_feature
1)..(7252)

;  OTHER INFORMATION: plasmid vector
;  PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: AX003206
;  DATABASE ENTRY DATE: 2000-08-24
;  RELEVANT RESIDUES: (1)..(7252)
US-09-238-356-27

  Query Match          36.9%;  Score 28.4;  DB 4;  Length 7252;
  Best Local Similarity  96.7%;  Pred. No. 0.23;
  Matches  29;  Conservative    0;  Mismatches    1;  Indels    0,

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | |||||||||||||||||||||||||||||
Db       2076 CGAGCATCATCCTCTGCATGGTCAGGTCAT 2105


RESULT 39
5169760-3
;Patent No. 5169760
;    APPLICANT: WILCON, EDWARD R.
;    TITLE OF INVENTION: METHOD, VECTORS AND HOST CELLS FOR
;THE CONTROL OF EXPRESSION OF HETEROLOGOUS GENES FROM LAC
;OPERATED PROMOTERS
;    NUMBER OF SEQUENCES: 3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/386,821
;      FILING DATE: 27-JUL-1989
;SEQ ID NO:3:
;      LENGTH: 7366
5169760-3

  Query Match          36.9%;  Score 28.4;  DB 6;  Length 7366;
  Best Local Similarity  96.7%;  Pred. No. 0.23;
  Matches  29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | |||||||||||||||||||||||||||||
Db       2246 CGAGCATCATCCTCTGCATGGTCAGGTCAT 2275


RESULT 40
US-09-238-356-28/c
; Sequence 28, Application US/09238356
; Patent No. 6312683
; GENERAL INFORMATION:
;  APPLICANT:  Kingsman, et al
;  TITLE OF INVENTION:  Retroviral Vectors
;  FILE REFERENCE:  674523-2006
;  CURRENT APPLICATION NUMBER: US/09/238,356
;  CURRENT FILING DATE:  1999-01-27
;  PRIOR APPLICATION NUMBER:  PCT/GB/03876
;  PRIOR FILING DATE:  1998-12-22
;  NUMBER OF SEQ ID NOS:  64
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 28
;  LENGTH: 7387
;  TYPE: DNA
;  ORGANISM: Artificial Sequence, primer
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(7387)
;  OTHER INFORMATION: plasmid vector
;  PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: AX003207
;  DATABASE ENTRY DATE: 2000-08-24
;  RELEVANT RESIDUES: (1)..(7387)
US-09-238-356-28

  Query Match          36.9%;  Score 28.4;  DB 4;  Length 7387;
  Best Local Similarity  96.7%;  Pred. No. 0.23;
  Matches  29;  Conservative    0;  Mismatches    1;  Indels    0;

Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
```

CGAGCATCATCCTCTGCATGGTCAGGTCAT 2571

ed: April 13, 2004, 20:16:05
secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

:leic search, using sw model

April 13, 2004, 17:09:09 ; Search time 217.5 Seconds
(without alignments)
1503.960 Million cell updates/sec

US-10-090-326-25
77
1 acttacncagcccagcanca.....ggagnacaacagccancagt 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

3373863 seqs, 2124099041 residues

hits satisfying chosen parameters: 6747726

.ength: 0
.ength: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

is the number of results predicted by chance to have a
ıter than or equal to the score of the result being printed,
'ived by analysis of the total score distribution.

SUMMARIES

Query Match %	Length	DB	ID	Description
94.8	77	6	ABQ82559	Abq82559 Tyrosinas
39.0	1059	2	AAT29076	Aat29076 Tyrosinas
39.0	1059	2	AAT29077	Aat29077 Tyrosinas
39.0	1059	8	ADB16200	Adb16200 Cleavase
39.0	1059	8	ADB16199	Adb16199 Cleavase
39.0	1587	2	AAT29078	Aat29078 Tyrosinas
39.0	1587	2	AAT29079	Aat29079 Tyrosinas
39.0	1587	4	AAF26897	Aaf26897 Human tyr
39.0	1587	8	ADB16202	Adb16202 Cleavase
39.0	1587	8	ADB16201	Adb16201 Cleavase
39.0	1590	5	AAF88041	Aaf88041 Human tyr
39.0	1886	1	AAN80637	Aan80637 Encodes h
39.0	1894	2	AAQ06552	Aaq06552 Clone BBT
39.0	1894	2	AAT01009	Aat01009 Human tyr
39.0	1894	2	AAT75678	Aat75678 Human SK2
39.0	1906	2	AAQ66548	Aaq66548 Human tyr
39.0	1910	2	AAT35901	Aat35901 Tyrosinas
39.0	1964	7	ABQ83843	Abq83843 Human tyr
39.0	1964	9	ADC09569	Adc09569 Tyrosinas
39.0	2041	4	AAS60958	Aas60958 Human can
39.0	2384	2	AAQ72871	Aaq72871 Human tyr
39.0	2384	2	AAT33316	Aat33316 Tyrosinas
39.0	2384	2	ABQ76197	Abq76197 Human tum
24	30	39.0	6408	7	ABX10643	Abx10643
25	29.2	37.9	8892	2	AAZ11453	Aaz11453
26	28.4	36.9	67	2	AAX87446	Aax87446
27	28.4	36.9	67	2	AAX87448	Aax87448
28	28.4	36.9	100	9	ADC60826	Adc60826
29	28.4	36.9	1542	6	ABA99646	Aba99646
30	28.4	36.9	3144	6	ABL61145	Abl61145
31	28.4	36.9	3150	4	AAH45081	Aah45081
32	28.4	36.9	3157	6	ABS70119	Abs70119
33	28.4	36.9	3157	7	ACC00558	Acc00558
34	28.4	36.9	3157	8	ACC71946	Acc71946
35	28.4	36.9	3279	6	AAD22647	Aad22647
36	28.4	36.9	3284	9	ADB89192	Adb89192
37	28.4	36.9	3284	9	ADC24539	Adc24539
38	28.4	36.9	3284	9	ADD67809	Add67809
39	28.4	36.9	3284	9	ADE10522	Ade10522
40	28.4	36.9	3284	9	ADE11444	Ade11444
41	28.4	36.9	3284	9	ADE12622	Ade12622
42	28.4	36.9	3284	9	ADE12385	Ade12385
43	28.4	36.9	3365	2	AAQ41287	Aaq41287
44	28.4	36.9	3396	8	ACF57390	Acf57390
45	28.4	36.9	3652	5	AAD04924	Aad04924

ALIGNMENTS

RESULT 1
ABQ82559
ID ABQ82559 standard; DNA; 77 BP.
XX
AC ABQ82559;
XX
DT 18-DEC-2002 (first entry)
XX
DE Tyrosinase mimic internal control probe SEQ ID NO:25.
XX
KW Human; CEA; carcinoembryonic antigen; adenocarcinoma; oesophagı
KW malignancy; probe; ss.
XX
OS Synthetic.
XX
PN WO200270751-A1.
XX
PD 12-SEP-2002.
XX
PF 04-MAR-2002; 2002WO-US006504.
XX
PR 02-MAR-2001; 2001US-0273277P.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Godfrey TE, Luketich JD, Raja S, Kelly LA, Finkelstein SD;
XX
DR WPI; 2002-732795/79.
XX
PT Multiplex PCR method for detecting malignancies, e.g. adenocarc
PT the esophagus comprises conducting a PCR amplification on a DNA
PT a PCR reaction mixture.
XX
PS Claim 47; Page 80; 141pp; English.
XX
CC The present invention describes a multiplex polymerase chain re
CC (PCR) (M1) comprising conducting PCR on a DNA sample in a react
CC mixture conducted in first and second amplification stages, eac
CC or more PCR cycles comprising denaturing, annealing and elongat
CC where the elongating step may be conducted at the same temperat
CC annealing step. The second amplification stage of (M1) is condu
CC different reaction conditions from that of the first amplificat
CC to modulate the relative rate of production of the first amplic
CC first primer set and a second amplicon by a second primer set d
CC first and second amplification stages. Also described: (1) an
CC oligonucleotide comprising 15-28 bp or its derivative; (2) intı

>stic; (3) rapid detection of a malignancy or of metastasised
inoma of the oesophagus; or (4) a cartridge for use in an
PCR system. (M1) is useful for detecting malignancies, e.g.
inoma of the oesophagus. (M1) eliminates contamination and
the time it takes to carry out a PCR reaction. The present
epresents an internal control probe, which is used in an
om the present invention

77 BP; 20 A; 22 C; 15 G; 16 T; 0 U; 4 Other;

94.8%; Score 73; DB 6; Length 77;
milarity 100.0%; Pred. No. 3.8e-18;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGA 60

NACAACAGCCANCAGT 77
|||||||||||||||
NACAACAGCCANCAGT 77

tandard; DNA; 1059 BP.

6 (first entry)

gene fragment.

t; mutation; cleavage; nuclease; cleavase; Thermus;
a; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
ccus; identification; detection; ss.

ns.

A1.

6.

5; 95WO-US014673.

4; 94US-00337164.
5; 95US-00402601.
5; 95US-00484956.
5; 95US-00520946.

IRD WAVE TECHNOLOGIES INC.

E, Lyamichev VI, Brow MAD, Oldenburg MC, Heisler LM;
live DM;

259862/26.

f nucleic acids to detect mutation(s) - allows detection esp.
53 gene, to identify strains of microorganisms and viruses.

; Page 268; 433pp; English.

f nucleic acids using an enzyme, especially a nuclease selected
roup consisting of Cleavase (RTM) BN enzyme, Thermus aquaticus
rase, Thermus thermophilus DNA polymerase, Escherichia coli
the Saccharomyces cerevisiae Rad1/Rad10 complex. The nucleic
rate is preferably an oligonucleotide containing a human p53
nce or alternatively, microbial gene sequences. Cleavage
re compared to the cleavage products of reference gene
The method is used for detecting mutation in the human p53
identifying strains of microorganisms, especially bacteria
rom the the group of members of the genera Campylobacter,
a, Mycobacterium, Salmonella, Shigella and Staphylococcus. The

CC method may also be used for the identification of viruses, esp
CC hepatitis C virus and simian immunodeficiency virus. The huma
CC gene (both wild type and mutant gene fragments) was used as a
CC sequence for the method. This sequence is a fragment of the t
CC gene (Exons 1-4)
XX
SQ Sequence 1059 BP; 298 A; 248 C; 227 G; 286 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 1059;
Best Local Similarity 67.9%; Pred. No. 0.33;
Matches 53; Conservative 0; Mismatches 24; Indels 1,

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | | || | | ||
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 535 AGTACAACAGCCATCAGT 552

RESULT 3
AAT29077
ID AAT29077 standard; DNA; 1059 BP.
XX
AC AAT29077;
XX
DT 02-DEC-1996 (first entry)
XX
DE Tyrosinase gene fragment (422 mutant).
XX
KW p53; mutant; mutation; cleavage; nuclease; cleavase; Thermus;
KW Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shig
KW Staphylococcus; identification; detection; ss.
XX
OS Homo sapiens.
XX
PN WO9615267-A1.
XX
PD 23-MAY-1996.
XX
PF 09-NOV-1995; 95WO-US014673.
XX
PR 09-NOV-1994; 94US-00337164.
PR 09-MAR-1995; 95US-00402601.
PR 07-JUN-1995; 95US-00484956.
PR 30-AUG-1995; 95US-00520946.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Dahlberg JE, Lyamichev VI, Brow MAD, Oldenburg MC, Heisler
PI Fors L, Olive DM;
XX
DR WPI; 1996-259862/26.
XX
PT Cleavage of nucleic acids to detect mutation(s) - allows detec
PT in human p53 gene, to identify strains of microorganisms and v
XX
PS Example 10; Page 268-269; 433pp; English.
XX
CC Cleavage of nucleic acids using an enzyme, especially a nuclea
CC from the group consisting of Cleavase (RTM) BN enzyme, Thermus
CC DNA polymerase, Thermus thermophilus DNA polymerase, Escherich
CC ExoIII and the Saccharomyces cerevisiae Rad1/Rad10 complex. Th
CC acid substrate is preferably an oligonucleotide containing a h
CC gene sequence or alternatively, microbial gene sequences. Clea
CC products are compared to the cleavage products of reference ge
CC sequences. The method is used for detecting mutation in the hu
CC gene; for identifying strains of microorganisms, especially ba
CC selected from the the group of members of the genera Campyloba
CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphyloc
CC method may also be used for the identification of viruses, esp
CC hepatitis C virus and simian immunodeficiency virus. The human

wild type and mutant gene fragments) was used as a test
or the method. This sequence is a fragment of the tyrosinase
s 1-4 of the 422 mutant)

059 BP; 299 A; 248 C; 226 G; 286 T; 0 U; 0 Other;

39.0%; Score 30; DB 2; Length 1059;
milarity 67.9%; Pred. No. 0.33;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 552

tandard; DNA; 1059 BP.

3 (first entry)

N DNA substrate #10.

lymerase; microorganism strain identification; bacteria;
ter; Escherichia; Mycobacterium; Salmonella; Shigella;
ccus; virus; hepatitis C virus; simian immunodeficiency virus;
ium tuberculosis; human.

ns.

38-A1.

3.

1; 2001US-00940925.

2; 92US-00986330.
3; 93US-00073384.
4; 94US-00254359.
4; 94US-00337164.
5; 95US-00402601.
5; 95US-00484956.
5; 95US-00520946.
7; 97US-00789079.
7; 97US-00802233.
); 2000US-00655378.

HLBERG J E.
OW M A D.
AMICHEV V I.

3, Brow MAD, Lyamichev VI;

615811/58.

:ion of strains of microorganisms, by treating nucleic acid
:ructure(s) derived from microorganisms with nuclease to form
:oducts(s) and detecting the product(s).

Page 131-132; 303pp; English.

.on relates to a method of detecting and identifying strains of
.sms by providing a nuclease and a nucleic acid substrate
sequences derived from microorganism(s), treating the nucleic
:ate to form cleavage structure(s) and reacting the nuclease
.eavage structures so that cleavage product(s) are produced.
is used for the identification of strains of microorganisms.

CC The microorganism comprises bacteria including Campylobacter,
CC Escherichia, Mycobacterium, Salmonella, Shigella or Staphyloc
CC virus comprising hepatitis C virus or simian immunodeficiency
CC Mycobacterium comprises strains of multi-drug resistant Mycob
CC tuberculosis. The method is less sensitive to size so that ent
CC rather than gene fragments, may be analysed. It facilitates th
CC internal standards for subsequent analysis and data comparison
CC increases the productivity of personnel and equipment. The pre
CC sequence represents a Cleavase BN substrate DNA.
XX
SQ Sequence 1059 BP; 299 A; 248 C; 226 G; 286 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 8; Length 1059;
Best Local Similarity 67.9%; Pred. No. 0.33;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
Db 535 AGTACAACAGCCATCAGT 552

RESULT 5
ADB16199
ID ADB16199 standard; DNA; 1059 BP.
XX
AC ADB16199;
XX
DT 20-NOV-2003 (first entry)
XX
DE Cleavase BN DNA substrate #9.
XX
KW ds; DNA polymerase; microorganism strain identification; bacte
KW Campylobacter; Escherichia; Mycobacterium; Salmonella; Shigell
KW Staphylococcus; virus; hepatitis C virus; simian immunodeficie
KW Mycobacterium tuberculosis; human.
XX
OS Homo sapiens.
XX
PN US2003054338-A1.
XX
PD 20-MAR-2003.
XX
PF 28-AUG-2001; 2001US-00940925.
XX
PR 07-DEC-1992; 92US-00986330.
PR 04-JUN-1993; 93US-00073384.
PR 06-JUN-1994; 94US-00254359.
PR 09-NOV-1994; 94US-00337164.
PR 09-MAR-1995; 95US-00402601.
PR 07-JUN-1995; 95US-00484956.
PR 30-AUG-1995; 95US-00520946.
PR 06-FEB-1997; 97US-00789079.
PR 19-FEB-1997; 97US-00802233.
PR 05-SEP-2000; 2000US-00655378.
XX
PA (DAHL/) DAHLBERG J E.
PA (BROW/) BROW M A D.
PA (LYAM/) LYAMICHEV V I.
XX
PI Dahlberg JE, Brow MAD, Lyamichev VI;
XX
DR WPI; 2003-615811/58.
XX
PT Identification of strains of microorganisms, by treating nuclei
PT cleavage structure(s) derived from microorganisms with nuclease
PT cleavage products(s) and detecting the product(s).
XX
PS Example 12; Page 130-131; 303pp; English.
XX

:ion relates to a method of detecting and identifying strains of
nisms by providing a nuclease and a nucleic acid substrate
g sequences derived from microorganism(s), treating the nucleic
:rate to form cleavage structure(s) and reacting the nuclease
:leavage structures so that cleavage product(s) are produced.
l is used for the identification of strains of microorganisms.
)rganism comprises bacteria including Campylobacter,
ia, Mycobacterium, Salmonella, Shigella or Staphylococcus or a
)rising hepatitis C virus or simian immunodeficiency virus. The
rium comprises strains of multi-drug resistant Mycobacterium
sis. The method is less sensitive to size so that entire genes,
an gene fragments, may be analysed. It facilitates the use of
standards for subsequent analysis and data comparison, and
the productivity of personnel and equipment. The present
represents a Cleavase BN substrate DNA.

1059 BP; 298 A; 248 C; 227 G; 286 T; 0 U; 0 Other;

39.0%; Score 30; DB 8; Length 1059;
imilarity 67.9%; Pred. No. 0.33;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| |||||| | | || | | || ||||||| |
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

AGNACAACAGCCANCAGT 77
| |||||||||| ||||
AGTACAACAGCCATCAGT 552

standard; DNA; 1587 BP.

)6 (first entry)

gene fragment.

it; mutation; cleavage; nuclease; cleavase; Thermus;
.a; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
)ccus; identification; detection; ss.

ins.

A1.

)6.

)5; 95WO-US014673.

)4; 94US-00337164.
)5; 95US-00402601.
)5; 95US-00484956.
5; 95US-00520946.

IRD WAVE TECHNOLOGIES INC.

E, Lyamichev VI, Brow MAD, Oldenburg MC, Heisler LM;
)live DM;

259862/26.

)f nucleic acids to detect mutation(s) - allows detection esp.
)53 gene, to identify strains of microorganisms and viruses.

; Page 268-269; 433pp; English.

)f nucleic acids using an enzyme, especially a nuclease selected
roup consisting of Cleavase (RTM) BN enzyme, Thermus aquaticus
rase, Thermus thermophilus DNA polymerase, Escherichia coli

CC ExoIII and the Saccharomyces cerevisiae Rad1/Rad10 complex. T
CC acid substrate is preferably an oligonucleotide containing a
CC gene sequence or alternatively, microbial gene sequences. Cle
CC products are compared to the cleavage products of reference g
CC sequences. The method is used for detecting mutation in the h
CC gene; for identifying strains of microorganisms, especially b
CC selected from the the group of members of the genera Campylob
CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphylo
CC method may also be used for the identification of viruses, es
CC hepatitis C virus and simian immunodeficiency virus. The huma
CC gene (both wild type and mutant gene fragments) was used as a
CC sequence for the method. This sequence is a fragment of the t
CC gene (Exons 1-5)
XX
SQ Sequence 1587 BP; 409 A; 386 C; 370 G; 422 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 1587;
Best Local Similarity 67.9%; Pred. No. 0.37;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | || | | || |
Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 842 AGTACAACAGCCATCAGT 859

RESULT 7
AAT29079
ID AAT29079 standard; DNA; 1587 BP.
XX
AC AAT29079;
XX
DT 02-DEC-1996 (first entry)
XX
DE Tyrosinase gene fragment (422 mutant).
XX
KW p53; mutant; mutation; cleavage; nuclease; cleavase; Thermus;
KW Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shig
KW Staphylococcus; identification; detection; ss.
XX
OS Homo sapiens.
XX
PN WO9615267-A1.
XX
PD 23-MAY-1996.
XX
PF 09-NOV-1995; 95WO-US014673.
XX
PR 09-NOV-1994; 94US-00337164.
PR 09-MAR-1995; 95US-00402601.
PR 07-JUN-1995; 95US-00484956.
PR 30-AUG-1995; 95US-00520946.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Dahlberg JE, Lyamichev VI, Brow MAD, Oldenburg MC, Heisler
PI Fors L, Olive DM;
XX
DR WPI; 1996-259862/26.
XX
PT Cleavage of nucleic acids to detect mutation(s) - allows detec
PT in human p53 gene, to identify strains of microorganisms and v
XX
PS Example 10; Page 270-271; 433pp; English.
XX
CC Cleavage of nucleic acids using an enzyme, especially a nuclea
CC from the group consisting of Cleavase (RTM) BN enzyme, Thermus
CC DNA polymerase, Thermus thermophilus DNA polymerase, Escherich
CC ExoIII and the Saccharomyces cerevisiae Rad1/Rad10 complex. Th
CC acid substrate is preferably an oligonucleotide containing a h

nce or alternatively, microbial gene sequences. Cleavage
re compared to the cleavage products of reference gene
The method is used for detecting mutation in the human p53
identifying strains of microorganisms, especially bacteria
rom the the group of members of the genera Campylobacter,
a, Mycobacterium, Salmonella, Shigella and Staphylococcus. The
also be used for the identification of viruses, especially
C virus and simian immunodeficiency virus. The human tyrosinase
wild type and mutant gene fragments) was used as a test
or the method. This sequence is a fragment of the tyrosinase
s 1-5 of the 422 mutant)'

587 BP; 410 A; 386 C; 369 G; 422 T; 0 U; 0 Other;

39.0%; Score 30; DB 2; Length 1587;
nilarity 67.9%; Pred. No. 0.37;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| |||||| | | || | | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 859

:andard; cDNA; 1587 BP.

l (first entry)

sinase nucleotide sequence SEQ ID NO:3.

)sinase; anti-gray hair agent; hair colour; hair follicle; ss.

1s.

)8-A.

).

); 99JP-00117006.

); 99JP-00117006.

:SEIDO CO LTD.

.41361/15.
)51344.

:he effect of an anti-gray hair agent.

Page 5-7; 11pp; Japanese.

: invention describes a method for separating the mRNA encoding
in a hair follicle for the assumption of hair colour and for
ın anti-gray hair agent. The method is useful for screening an
iair agent and judging the effect of it. The present sequence
ian tyrosinase which is given in the exemplification of the
'ention

387 BP; 408 A; 387 C; 370 G; 422 T; 0 U; 0 Other;

39.0%; Score 30; DB 4; Length 1587;
ilarity 67.9%; Pred. No. 0.37;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

:TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||| |||||||||| |||||| | | || | | | |||||||

Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 842 AGTACAACAGCCATCAGT 859

RESULT 9
ADB16202
ID ADB16202 standard; DNA; 1587 BP.
XX
AC ADB16202;
XX
DT 20-NOV-2003 (first entry)
XX
DE Cleavase BN DNA substrate #12.
XX
KW ds; DNA polymerase; microorganism strain identification; bacte
KW Campylobacter; Escherichia; Mycobacterium; Salmonella; Shigell
KW Staphylococcus; virus; hepatitis C virus; simian immunodeficie
KW Mycobacterium tuberculosis; human.
XX
OS Homo sapiens.
XX
PN US2003054338-A1.
XX
PD 20-MAR-2003.
XX
PF 28-AUG-2001; 2001US-00940925.
XX
PR 07-DEC-1992; 92US-00986330.
PR 04-JUN-1993; 93US-00073384.
PR 06-JUN-1994; 94US-00254359.
PR 09-NOV-1994; 94US-00337164.
PR 09-MAR-1995; 95US-00402601.
PR 07-JUN-1995; 95US-00484956.
PR 30-AUG-1995; 95US-00520946.
PR 06-FEB-1997; 97US-00789079.
PR 19-FEB-1997; 97US-00802233.
PR 05-SEP-2000; 2000US-00655378.
XX
PA (DAHL/) DAHLBERG J E.
PA (BROW/) BROW M A D.
PA (LYAM/) LYAMICHEV V I.
XX
PI Dahlberg JE, Brow MAD, Lyamichev VI;
XX
DR WPI; 2003-615811/58.
XX
PT Identification of strains of microorganisms, by treating nucle:
PT cleavage structure(s) derived from microorganisms with nuclease
PT cleavage products(s) and detecting the product(s).
XX
PS Example 12; Page 133; 303pp; English.
XX
CC The invention relates to a method of detecting and identifying
CC microorganisms by providing a nuclease and a nucleic acid subst
CC containing sequences derived from microorganism(s), treating th
CC acid substrate to form cleavage structure(s) and reacting the r
CC with the cleavage structures so that cleavage product(s) are pr
CC The method is used for the identification of strains of microor
CC The microorganism comprises bacteria including Campylobacter,
CC Escherichia, Mycobacterium, Salmonella, Shigella or Staphylococ
CC virus comprising hepatitis C virus or simian immunodeficiency v
CC Mycobacterium comprises strains of multi-drug resistant Mycobac
CC tuberculosis. The method is less sensitive to size so that enti
CC rather than gene fragments, may be analysed. It facilitates the
CC internal standards for subsequent analysis and data comparison,
CC increases the productivity of personnel and equipment. The pres
CC sequence represents a Cleavase BN substrate DNA.
XX
SQ Sequence 1587 BP; 410 A; 386 C; 369 G; 422 T; 0 U; 0 Other;

39.0%; Score 30; DB 8; Length 1587;
milarity 67.9%; Pred. No. 0.37;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 859

tandard; DNA; 1587 BP.

3 (first entry)

N DNA substrate #11.

lymerase; microorganism strain identification; bacteria;
ter; Escherichia; Mycobacterium; Salmonella; Shigella;
ccus; virus; hepatitis C virus; simian immunodeficiency virus;
ium tuberculosis; human.

ns.

38-A1.

3.

1; 2001US-00940925.

2; 92US-00986330.
3; 93US-00073384.
4; 94US-00254359.
4; 94US-00337164.
5; 95US-00402601.
5; 95US-00484956.
5; 95US-00520946.
7; 97US-00789079.
7; 97US-00802233.
0; 2000US-00655378.

HLBERG J E.
OW M A D.
AMICHEV V I.

E, Brow MAD, Lyamichev VI;

615811/58.

ition of strains of microorganisms, by treating nucleic acid
structure(s) derived from microorganisms with nuclease to form
oducts(s) and detecting the product(s).

; Page 132; 303pp; English.

ion relates to a method of detecting and identifying strains of
isms by providing a nuclease and a nucleic acid substrate
sequences derived from microorganism(s), treating the nucleic
rate to form cleavage structure(s) and reacting the nuclease
leavage structures so that cleavage product(s) are produced.
is used for the identification of strains of microorganisms.
organism comprises bacteria including Campylobacter,
ia, Mycobacterium, Salmonella, Shigella or Staphylococcus or a
prising hepatitis C virus or simian immunodeficiency virus. The
rium comprises strains of multi-drug resistant Mycobacterium
sis. The method is less sensitive to size so that entire genes,
an gene fragments, may be analysed. It facilitates the use of

CC internal standards for subsequent analysis and data comparison
CC increases the productivity of personnel and equipment. The pre
CC sequence represents a Cleavase BN substrate DNA.
XX
SQ Sequence 1587 BP; 409 A; 386 C; 370 G; 422 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 8; Length 1587;
Best Local Similarity 67.9%; Pred. No. 0.37;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
Db 782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
Db 842 AGTACAACAGCCATCAGT 859

RESULT 11
AAF88041
ID AAF88041 standard; cDNA; 1590 BP.
XX
AC AAF88041;
XX
DT 16-JUL-2001 (first entry)
XX
DE Human tyrosinase encoding cDNA.
XX
KW Listeria; expression vector; tumor-associated antigen; Trp 1;
KW MelanA/MART-1; cytostatic; attenuated; immunotherapy; malignan
KW pigmented tumor; malignant schwannoma; vaccination; tyrosinase
KW antigen-presenting cell; ss.
XX
OS Homo sapiens.
XX
PN WO200127295-A1.
XX
PD 19-APR-2001.
XX
PF 13-OCT-2000; 2000WO-DE003629.
XX
PR 14-OCT-1999; 99DE-01049594.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Schadendorf D, Paschen A, Chakraborty T, Domann E;
XX
DR WPI; 2001-282041/29.
DR P-PSDB; AAB86039.
XX
PT Listeria expression vector for immunotherapy, particularly of
PT melanoma, comprises a DNA sequence encoding tumor-associated a
XX
PS Disclosure; Fig 1; 41pp; German.
XX
CC This invention describes a novel Listeria expression vector (
CC immunotherapy which comprises a promoter (P), functional in Li
CC operably linked to a DNA sequence (I) encoding one of the tumc
CC associated antigens (II) human tyrosinase, Trp 1 or 2, or Mela
CC The products of the invention have cytostatic activity. Recomb
CC attenuated Listeria containing (A) are useful for immunotherap
CC (prophylactic, adjuvant or therapeutic), specifically of malig
CC melanoma (but also other pigmented tumors such as malignant sc
CC particularly as a replacement for radiotherapy. Using attenuat
CC as carrier for (A) provides a simple way of vaccination, since
CC presenting cells acquire tumor-associated antigens by natural
CC eliminating the need for labor-intensive ex vivo modification
CC autologous cells. This sequence encodes the human tyrosinase p
CC described in the method of the invention
XX
SQ Sequence 1590 BP; 411 A; 386 C; 370 G; 423 T; 0 U; 0 Other;

39.0%; Score 30; DB 5; Length 1590;
.milarity 67.9%; Pred. No. 0.37;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 859

tandard; cDNA; 1886 BP.

0 (first entry)

man tyrosinase.

sinase; albinism; melanoma; ss.

ns.

Location/Qualifiers
7. .1689
/*tag= a
e 7. .42
/*tag= b
al 1864. .1869
/*tag= c

A.

3.

7; 87WO-US002561.

5; 86US-00915753.

THRIE FOUND MED.

105497/15.
80624.

or human tyrosinase - used for prodn. of tyrosinase and human
id as probe for genetic analysis of albinism and melanoma.

Page ?; 36pp; English.

11 cDNA library of normal human melanocytes was screened with
directed against purified hamster tyrosinase. 13 of the 16
h were isolated were found to cross-react with each other.
gous to a representative cDNA lambda mel 34 (the longest cDNA
expressed specifically in melanocytes, detecting a 2.4kb mRNA
human melanocytes. The lambda mel 34 clone encoded human
The cDNA can be fused to an expression vector for prodn. of
of pure tyrosinase. It can also be used as a probe for
lysis of human albinism, melanotic and amelanotic melanoma and
al diagnosis of albinism

86 BP; 500 A; 446 C; 426 G; 514 T; 0 U; 0 Other;

39.0%; Score 30; DB 1; Length 1886;
ilarity 67.9%; Pred. No. 0.39;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
Db 785 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGGCG

Qy 60 AGNACAACAGCCANCAGT 77
Db 845 AGTACAACAGCCATCAGT 862

RESULT 13
AAQ06552
ID AAQ06552 standard; DNA; 1894 BP.
XX
AC AAQ06552;
XX
DT 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 15-MAR-1991 (first entry)
XX
DE Clone BBTY-1 encoding human tyrosinase.
XX
KW Tyrosinase; non-melanocytic cells; melanin; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1. .57
FT /*tag= a
FT mat_peptide 58. .1593
FT /*tag= b
FT /product= "mature tyrosinase"
XX
PN WO9012869-A.
XX
PD 01-NOV-1990.
XX
PF 26-APR-1989; 89US-00343960.
XX
PR 26-APR-1989; 89US-00343960.
XX
PA (SLOK) SLOAN KETTERING INST CANCER.
XX
PI Bouchard B, Hougton A;
XX
DR WPI; 1990-348472/46.
DR P-PSDB; AAR07071.
XX
PT Non-melanocyte eukaryotic cell - expressing biologically activ
PT tyrosinase and used in studying tyrosine synthesis.
XX
PS Disclosure; Fig 1; 49pp; English.
XX
CC This cDNA clone was isolated from a lambda gt10 library derive
CC pigmented human melanoma cell line SK-MEL-19 using the probe o
CC Following processing of the prod. a polypeptide of mol. wt. ca
CC obtd. which was found to be biologically active human tyrosina
CC constitutively expressed using a non-melanocytic, eukaryotic c
CC fibroblast cell. The tyrosinase enzyme catalyses the prodn. of
CC See also AAQ06551. (Updated on 09-JAN-2003 to add missing OS f
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1894 BP; 504 A; 447 C; 425 G; 518 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 1894;
Best Local Similarity 67.9%; Pred. No. 0.4;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATT
Db 788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGAT

Qy 60 AGNACAACAGCCANCAGT 77
Db 848 AGTACAACAGCCATCAGT 865

tandard; DNA; 1894 BP.

6 (first entry)

sinase acting as a tumor rejection antigen precursor.

ction antigen precuror; tyrosinase; enzyme; tumor therapy; ss.

ns.

Location/Qualifiers
e 1. .1593
/*tag= a

A1.

5.

5; 95WO-US001990.

4; 94US-00203054.
4; 94US-00233305.

DWIG INST CANCER RES.

Van Pel A, Brichard V, Boon-Falleur T;

311542/40.
R79493.

T cell specific for MHC complexes - used to treat individual(s)
lar abnormality, e.g. tumour(s).

; Page 30-33; 57pp; English.

ncodes human tyrosinase which may act as a tumor rejection
ecursor (TRAP) and be processed to form a peptide tumor
antigen which is presented on the surface of the cell, in
n with HLA-A2, thereby stimulating lysis by cytotoxic T-
clones

894 BP; 504 A; 447 C; 425 G; 518 T; 0 U; 0 Other;

39.0%; Score 30; DB 2; Length 1894;
milarity 67.9%; Pred. No. 0.4;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 865
```

tandard; cDNA; 1894 BP.

7 (first entry)

-MEL; tyrosinase cDNA.

tumour rejection antigen; TRAP; human leukocyte antigen;

KW HLA-B44; SK29-MEL; major histocompatibility complex; MHC;
KW cellular abnormality; melanoma; autoimmune disease; gene thera
KW diagnosis; ss.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .1596
FT		/*tag= a
FT		/transl_except= (pos:100. .102, aa:Gly)
FT		/transl_except= (pos:541. .543, aa:Ile)
FT		/transl_except= (pos:580. .582, aa:Tyr)
FT		/transl_except= (pos:667. .669, aa:Gly)
FT		/transl_except= (pos:691. .693, aa:Gly)
FT		/transl_except= (pos:754. .756, aa:Gly)
FT		/transl_except= (pos:1210. .1212, aa:Arg)
FT		/transl_except= (pos:1411. .1413, aa:Gly)
FT	sig_peptide	1. .54
FT		/*tag= b
FT	mat_peptide	55. .1593
FT		/*tag= c

XX
PN WO9726535-A1.
XX
PD 24-JUL-1997.
XX
PF 14-JAN-1997; 97WO-US000834.
XX
PR 17-JAN-1996; 96US-00587391.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lethe B, Brichard V, Van Pel A, Boon-Falleur T, Wolfel T;
XX
DR WPI; 1997-385477/35.
DR P-PSDB; AAW22083.
XX
PT Complex of human leukocyte antigen and tyrosinase peptide - us
PT treating cellular abnormalities, e.g. tumours.
XX
PS Example 5; Page 34-37; 64pp; English.
XX
CC This cDNA sequence codes for human tyrosinase (AAW22083). It w
CC from cell line SK29-MEL.1, a subclone of the melanoma cell lin
CC following a screen for clones that stimulated cytolytic T cell
CC 210/9 and CTL IVSB. Human tyrosinase acts as a tumour rejectio
CC precursor and is processed to form a peptide tumour rejection
CC (see AAW22082) that is presented on the surface of a cell in c
CC with HLA-A2, thereby stimulating lysis by CTL clones. Methods
CC for identifying individuals having abnormal cells that present
CC of the tumour rejection antigen and HLA-B44, for treating such
CC individuals (e.g. by administering vectors that code for human
CC and an HLA-B44 molecule), for isolating cytolytic T cells spec
CC the complex of HLA-B44 and the peptide, and for identifying ab
CC cells. The methods can be used to treat tumours, such as melan
CC autoimmune disorders, etc
XX
SQ Sequence 1894 BP; 504 A; 447 C; 425 G; 518 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 1894;
Best Local Similarity 67.9%; Pred. No. 0.4;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT'
              ||||||  |||||||||| ||||||       |   | |  ||  |   | |
Db        788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA'

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        848 AGTACAACAGCCATCAGT 865
```

RESULT 16

```
:andard; DNA; 1906 BP.


3  (revised)
5  (first entry)

sinase DNA.

; enzyme; tumor rejection antigen precursor; HLA;
)cyte antigen; ss.

1s.

      Location/Qualifiers
      8. .1611
      /*tag=  a

\1.

1.

3;    93WO-US012200.

2;    92US-00994928.
3;    93US-00054714.
3;    93US-00081673.

)WIG INST CANCER RES.

ir T,  Brichard V,  Van Pel A,  De Plaen E,  Coulie P;
  Wolfel T,  Lethe B;

234342/28.
256309.

 the identification of HLA complexes - used for the detection
. cells.

 Page 19; 31pp; English.

icodes a normally occurring tyrosinase which may act as a tumor
intigen precursor and be processed to form a peptide tumor
intigen (AAR56310) which is presented on the surface of a cell,
:ion with HLA-A2, thereby stimulating lysis of cytotoxic T-
 clones. (Updated on 25-MAR-2003 to correct PN field.)

'06 BP; 503 A; 450 C; 431 G; 522 T; 0 U; 0 Other;

          39.0%;  Score 30;  DB 2;  Length 1906;
ilarity   67.9%;  Pred. No. 0.4;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
TTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 865

NACAACAGCCANCAGT 77
  |||||||||| ||||
TACAACAGCCATCAGT 883



:andard; cDNA; 1910 BP.


;  (first entry)

 melanoma antigen DNA.

KW   Melanoma; tyrosinase; immunogen; vaccine; cancer; immunotherap
KW   transgenic animal; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             29. .1618
FT                   /*tag=  a
XX
PN   WO9621734-A2.
XX
PD   18-JUL-1996.
XX
PF   11-JAN-1996;   96WO-US000473.
XX
PR   10-JAN-1995;   95US-00370909.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Robbins PF,  Rosenberg SA;
XX
DR   WPI; 1996-342287/34.
DR   P-PSDB; AAW03306.
XX
PT   Nucleic acid sequence encoding p15 melanoma antigen - and immu
PT   peptide(s) derived from it, useful for diagnosis, prevention o
PT   of melanoma.
XX
PS   Disclosure; Page 76-77; 98pp; English.
XX
CC   A DNA fragment (AAT35901) codes for tyrosinase (AAW03306), an
CC   involved in melanin synthesis. The tyrosinase was recognised b
CC   restricted tumour infiltrating lymphocytes from a melanoma pat
CC   DNA fragment can be used for recombinant prodn. of tyrosinase
CC   immunogenic peptides (see also AAW03304-05) for use as melanom
XX
SQ   Sequence 1910 BP; 502 A; 451 C; 433 G; 524 T; 0 U; 0 Other;

  Query Match             39.0%;  Score 30;  DB 2;  Length 1910;
  Best Local Similarity   67.9%;  Pred. No. 0.4;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        810 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        870 AGTACAACAGCCATCAGT 887


RESULT 18
ABQ83843
ID   ABQ83843 standard; cDNA; 1964 BP.
XX
AC   ABQ83843;
XX
DT   03-FEB-2003  (first entry)
XX
DE   Human tyrosinase encoding cDNA SEQ ID NO:5.
XX
KW   Human; epitope; vaccine; immunotherapeutic; cytostatic; immunoc
KW   T cell; enzyme; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200281646-A2.
XX
PD   17-OCT-2002.
XX
PF   04-APR-2002; 2002WO-US011101.
XX
PR   06-APR-2001; 2001US-0282211P.
```

1; 2001US-0337017P.
2; 2002US-0363210P.

L IMMUNOTHERAPIES CORP.

, Diamond DC, Liu L, Xie Z;

067518/06.
P74129.

opes useful as vaccines, comprises peptides or nucleic acid
he peptides, that are useful epitopes of target-associated

age 146-147; 352pp; English.

t invention describes an isolated epitope (I) and an epitope
lso described is a vaccine or immunotherapeutic composition
ising (I). (I) has cytostatic activity. VC is useful for
n animal, by administering to an animal the vaccine or
apeutic composition. VC is also useful for evaluating
city of a vaccine or immunotherapeutic composition, by
ing VC to an HLA-transgenic animal and evaluating
city based on a characteristic of the animal, or by in vitro
imulation of a T cell and evaluating immunogenicity. (I) is
determining specific T cell frequency, by contacting T cells
-peptide complex, and further comprises ELISPOT analysis,
ilution analysis, flow cytometry, in situ hybridisation and/or
chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
epresent sequences used in the exemplification of the present

964 BP; 520 A; 462 C; 445 G; 537 T; 0 U; 0 Other;

39.0%; Score 30; DB 7; Length 1964;
nilarity 67.9%; Pred. No. 0.4;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||  ||||||||||  ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 923

GNACAACAGCCANCAGT 77
|  |||||||||  ||||
GTACAACAGCCATCAGT 941
```

tandard; cDNA; 1964 BP.

3 (first entry)

cDNA #SEQ ID 5.

mmunological; vaccine;
ocompatibility complex class I; MHC class I; cancer;
on; ss.

ed.

37-A2.

3.

2; 2002WO-US010189.

1; 2001US-0282211P.
1; 2001US-0337017P.
2; 2002US-0363210P.

PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard JJL, Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-248010/24.
XX
PT Epitope having high affinity for major histocompatibility comp
PT useful for treating an animal, evaluating immunogenicity of a
PT therapeutic composition and for diagnosing a disease.
XX
PS Claim 1; SEQ ID NO 5; 239pp; English.
XX
CC The invention relates to an isolated epitope polypeptide that
CC affinity for major histocompatibility complex (MHC) class I, a
CC epitope cluster comprising the polypeptide. Also disclosed is
CC or immunotherapeutic composition containing an epitope of the
CC Compositions of the invention may be used in the treatment of
CC method can be combined with a radiation therapy, chemotherapy,
CC biochemotherapy or surgery. The composition is also useful for
CC immunogenicity of a vaccine or immunotherapeutic compound. Mul
CC -peptide complexes of the invention are useful for determining
CC cell frequency. This method is useful for evaluating immunolog
CC response, by performing the method prior to and subsequent to
CC immunisation step. Compositions of the invention are useful fo
CC diagnosing a disease. The current sequence represents an epito
CC invention with high affinity for MHC class I.
XX
SQ Sequence 1964 BP; 520 A; 462 C; 445 G; 537 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 9; Length 1964;
Best Local Similarity 67.9%; Pred. No. 0.4;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

```
Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT 
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        864 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        924 AGTACAACAGCCATCAGT 941
```

RESULT 20
AAS60958
ID AAS60958 standard; cDNA; 2041 BP.
XX
AC AAS60958;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human cancer agent-resistance marker #617.
XX
KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinom
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.
XX
PN WO200179556-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US012132.
XX
PR 14-APR-2000; 2000US-0197538P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
DR WPI; 2001-602933/68.
XX

:ic acid, used as a marker to determine the effectiveness of
, to treat cancer cell growth in individuals.

.ge 460-461; 527pp; English.

on relates to 1046 novel nucleic acids which are used as
determining the sensitivity of a cancer cell to the
agent TAXOL. Cancer cells can be treated with TAXOL when they
.o express one of the 242 sensitivity markers or the cells are
.o express one of the 804 resistance markers. The methods can
determine the effectiveness of TAXOL in the treatment of
. growth in an individual. The markers can be used as targets
.ng anti-cancer agents such as chemotherapeutic compounds. The
ı also be used as targets in developing treatments for cancer,
.y those cancers which display resistance to agents and exhibit
of the markers. The anticancer agents developed by the novel
be used to treat cancer. Probes based on the markers can be
:ect transcripts or genomic sequences corresponding to the
ı the identification of cells or tissues which mis-express the
ıncers which may be targeted include carcinoma (e.g. squamous
ıoma), sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic
lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's
l tumours (e.g. glioma). The present sequence is one of the
cancer cell markers

)41 BP; 564 A; 467 C; 455 G; 548 T; 0 U; 7 Other;

39.0%; Score 30; DB 4; Length 2041;
ilarity 67.9%; Pred. No. 0.41;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
TTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 966

NACAACAGCCANCAGT 77
TACAACAGCCATCAGT 984

andard; cDNA; 2384 BP.

(revised)
(first entry)

inase gene.

sinase; compound; hair; follicle; liposome; lipophilic;
stratum corneum; melanin; hair dye; glycoprotein; alopecia;
y; transforming growth factor; growth stimulant; aromatase;
A; macromolecule; polymer; ss.

s.

Location/Qualifiers
503. .2092
/*tag= a
/product= "human tyrosinase"

.1.

; 94WO-US003634.

; 93US-00041553.
; 94US-00181471.

ICANCER INC.

XX
PI Li L, Lishko VK;
XX
DR WPI; 1994-332816/41.
DR P-PSDB; AAR63623.
XX
PT Lipsomes for delivering protein, nucleic acid etc. to hair fol
PT e.g. to restore hair colour, prevent hair loss during chemothe
PT stimulate hair growth etc.
XX
PS Claim 6; Page 67-70; 100pp; English.
XX
CC The nucleotide sequence of a human tyrosinase gene. This is an
CC a compound which can be delivered to hair follicles via a novel
CC composition. The liposomes are comprised of a lipophilic or li
CC compound which will selectively target the hair follicle (via t
CC corneum) without damaging or unwanted effects on cells outside
CC follicle. Compounds e.g. tyrosinase (or the DNA encoding it), r
CC hair dyes, can be delivered to the hair follicle to restore ha
CC or condition. Other compounds targeted at hair follicles can in
CC glycoprotein (AAQ72872) (for treatment of chemotherapy-induced
CC human transforming growth factor alpha (AAQ72873)(for reversal
CC hair); hair growth stimulants (e.g. cyclosporin A or aromatase)
CC antisense sequences. The methods allows compounds (e.g. macrom
CC polymers), which would not normally reach the hair follicles, t
CC delivered to these target areas. (Updated on 25-MAR-2003 to cor
CC field.)
XX
SQ Sequence 2384 BP; 652 A; 534 C; 519 G; 679 T; 0 U; 0 Other;

Query Match 39.0%; Score 30; DB 2; Length 2384;
Best Local Similarity 67.9%; Pred. No. 0.43;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATT
Db 1284 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGAT

Qy 60 AGNACAACAGCCANCAGT 77
Db 1344 AGTACAACAGCCATCAGT 1361

RESULT 22
AAT33316
ID AAT33316 standard; cDNA to mRNA; 2384 BP.
XX
AC AAT33316;
XX
DT 18-OCT-1996 (first entry)
XX
DE Tyrosinase mRNA.
XX
KW Human; tyrosinase; detection; primer; reverse transcription; an
KW transcription region; detection; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 503. .2092
FT /*tag= a
FT /product= "Tyrosinase"
FT primer_bind 790. .810
FT /*tag= b
FT /note= "Binds primer given in AAT33323"
FT primer_bind 881. .901
FT /*tag= c
FT /note= "Binds primer given in AAT33321"
FT primer_bind 1297. .1317
FT /*tag= d
FT /note= "Binds primer given in AAT33317"
FT primer_bind 1327. .1347
FT /*tag= e

```
      /note= "Binds primer given in AAT33319"
d     complement(1521. .1541)
      /*tag=  f
      /note= "Binds primer given in AAT33318"
d     complement(1600. .1620)
      /*tag=  g
      /note= "Binds primer given in AAT33322"
d     complement(1650. .1670)
      /*tag=  h
      /note= "Binds primer given in AAT33320"
d     complement(1673. .1693)
      /*tag=  i
      /note= "Binds primer given in AAT33324"

-A.

6.

4;   94JP-00288041.

4;   94JP-00288041.

LA CHEM IND INC.

316329/32.
W00184.

of tyrosinase mRNA - by amplification of tyrosinase mRNA allows
etection of trace amounts of RNA.

; Page 7-9; 10pp; Japanese.

nce encodes human tyrosinase. This sequence was detected by the
the invention using the primers given in AAT33317-24. The
prises synthesising tyrosinase cDNA from a sample by reverse
ion and then amplifying the cDNA using primers specific for, or
ary to, sequences within the transcription region of the
 gene. Amplification using these primers produces a competitive
 can also be amplified. This method can be used to specifically
ce amounts of human tyrosinase mRNA

384 BP; 652 A; 533 C; 520 G; 679 T; 0 U; 0 Other;

          39.0%;  Score 30;  DB 2;  Length 2384;
nilarity   67.9%;  Pred. No. 0.43;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 1343

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 1361



tandard; DNA; 2384 BP.



2  (first entry)

ur antigen tyronsinase DNA.

igen; human; vaccine; cellular immune response; immunogen;
mour; tyrosinase; ds.

ns.

B1.
```

```
PD   11-SEP-2001.
XX
PF   06-APR-1998;   98US-00056105.
XX
PR   10-APR-1997;   97US-0043467P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Kipps TJ,  Wu Y;
XX
DR   WPI; 1998-583198/49.
XX
PT   Generating cellular immune response in patient to target prote
PT   comprises introducing vector with nucleotide sequence encoding
PT   comprising protein processing signal into cell of patient.
XX
PS   Disclosure; Col 19-22; 61pp; English.
XX
CC   This invention describes a novel method for generating a cellu
CC   response in a patient to a target protein or its fragment. The
CC   involves introducing a vector containing a nucleotide sequence
CC   chimeric immunogen comprising a protein processing signal and
CC   protein or its fragment. The immunogen is produced by the cell
CC   processed so that the target protein or its fragment is presen
CC   patients immune system and a cellular immune response is initi
CC   method and vectors can be used as a form of vaccination and co
CC   to generate a cellular immune response in patients to, e.g. ca
CC   tumours. The cellular immune response is the predominant immun
CC   in the patient. This sequence represents a DNA fragment which
CC   human tumour antigen tyrosinase described in the method of the
CC   Note: The information in this spec has been previously disclos
CC   WO199845444 however this spec contained no sequence informatio
XX
SQ   Sequence 2384 BP; 652 A; 533 C; 520 G; 679 T; 0 U; 0 Other;

  Query Match            39.0%;  Score 30;  DB 2;  Length 2384;
  Best Local Similarity  67.9%;  Pred. No. 0.43;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db       1284 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || ||||||||| ||||
Db       1344 AGTACAACAGCCATCAGT 1361


RESULT 24
ABX10643
ID   ABX10643 standard; DNA; 6408 BP.
XX
AC   ABX10643;
XX
DT   22-APR-2003  (first entry)
XX
DE   Vector, htyr-pING, containing human tyrosinase gene.
XX
KW   Human; gene; ds; melanoma; xenogeneic differentiation antigen;
KW   tyrosinase; canine; gene therapy; dog; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2002150589-A1.
XX
PD   17-OCT-2002.
XX
PF   27-NOV-2001; 2001US-00996128.
XX
PR   10-DEC-1996;   96US-0032535P.
PR   18-FEB-1997;   97US-0036419P.
PR   10-DEC-1997;   97WO-US022669.
PR   21-MAY-1999;   99US-00308697.
```

); 2000US-0180651P.
); 2000US-00627694.

JGHTON A N.
:GMAN P J.
:CHOK J D.

I, Bergman PJ, Wolchok JD;

.82484/18.

:lanoma in a mammalian subject comprises administering to the
immunological amount of a xenogeneic differentiation antigen
: type as a differentiation antigen expressed by melanoma cells
ect.

'age 5-8; 15pp; English.

.on discloses a method for treating melanoma in a mammalian
.ch comprises administering to the subject an immunological
: xenogeneic differentiation antigen (DA) of the same type as a
:d by melanoma cells of the subject. Also disclosed are pING
:taining a sequence encoding the human or mouse tyrosinase
methods and xenogeneic DA are useful for treating canine
:elanoma (gene therapy) in dog suffering from the disease by
.ng an immunological amount of the xenogeneic DA and for other
:e sequence presented is the vector, htyr-pING, containing the
:inase gene

08 BP; 1621 A; 1583 C; 1509 G; 1695 T; 0 U; 0 Other;

39.0%; Score 30; DB 7; Length 6408;
ilarity 67.9%; Pred. No. 0.6;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

TTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||| |||||||||| |||||| | | | || | | || |||||||
TTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 5356

NACAACAGCCANCAGT 77
|||||||||| ||||
TACAACAGCCATCAGT 5374

andard; DNA; 8892 BP.

(first entry)

vector pEICUT-Z sequence.

vector; functional splice donor site; hybrid viral vector;
splice acceptor site; in vivo gene delivery; therapeutic;
vector; modified hematopoietic stem cell; MHSC; tumour; MLV;
ypoxia response element; HRE; hypoxia; promoter; EIAV; ss.

2.

.

; 98WO-GB002885.

; 97GB-00020216.
; 97GB-00020465.

'ORD BIOMEDICA UK LTD.

inley KM, Bebbington C, Naylor S;

XX
DR WPI; 1999-263482/22.
XX
PT New retroviral vectors, for, e.g. delivering nucleotide sequen
PT solid tumor sites.
XX
PS Example 2 (page 172); Fig 11 (Page 12-13/35); 288pp; English.
XX
CC The invention relates to a retroviral vector (RVV) comprising
CC functional splice donor site (FSDS) and a functional splice ac
CC (FSAS) where: (i) the FSDS and the FSAS flank a first nucleoti
CC of interest (NOI); (ii) the FSDS is upstream of the FSAS; (iii
CC is derived from a retroviral pro-vector; (iv) the retroviral p
CC comprises a first nucleotide sequence (NS) capable of yielding
CC and a second NS capable of yielding the FSAS; and (v) the firs
CC downstream of the second NS, such that the RVV is formed as a
CC reverse transcription of the retroviral pro-vector. A hybrid v
CC (VV) system for in vivo gene delivery, which system comprises
CC VV which encodes a secondary VV, the primary vector capable of
CC a first target cell and of expressing the secondary VV, which
CC vector is capable of transducing a secondary target cell, wher
CC primary vector is obtainable from or is based on a adenoviral
CC the secondary VV is obtainable from or is based on a RVV prefe
CC lentiviral vector (LVV) is also provided. The systems can be u
CC delivering NOIs to one or more target sites. The NOIs may enco
CC therapeutic or diagnostic agents. The methods are used particu
CC producing modified hematopoietic stem cells (MHSCs) to deliver
CC sites such as solid tumours which are characterised by ischemi
CC hypoxia or low glucose concentration. The system permits the s
CC expression of NOIs in targeted cells, e.g. rapidly dividing ce
CC present sequence represents a EIAV lentiviral version of the v
CC (Intron Created Upon Transduction). The present plasmid is pEI
XX
SQ Sequence 8892 BP; 2116 A; 2233 C; 2345 G; 2198 T; 0 U; 0 Other;

Query Match 37.9%; Score 29.2; DB 2; Length 8892;
Best Local Similarity 91.2%; Pred. No. 1.4;
Matches 31; Conservative 0; Mismatches 3; Indels 0;

Qy 19 CATTCTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| | | ||||||||||||||||||||||||||||||
Db 2540 CGTACGAGCATCATCCTCTGCATGGTCAGGTCAT 2573

RESULT 26
AAX87446
ID AAX87446 standard; DNA; 67 BP.
XX
AC AAX87446;
XX
DT 08-OCT-1999 (first entry)
XX
DE Plasmid pJP5603 PCR primer b-left.
XX
KW ET cloning, recE; recT; Escherichia coli; PCR; primer; pJP5603;
KW homologous recombination; ss.
XX
OS Synthetic.
XX
PN WO9929837-A2.
XX
PD 17-JUN-1999.
XX
PF 07-DEC-1998; 98WO-EP007945.
XX
PR 05-DEC-1997; 97EP-00121462.
PR 05-OCT-1998; 98EP-00118756.
XX
PA (EMBL-) EMBL EURO LAB MOLEKULARBIOLOGIE.
XX
PI Stewart F, Zhang Y, Buchholz F;
XX
DR WPI; 1999-457893/38.

or cloning DNA using a recE and recT homologous recombinant

: Page 137; 139pp; English.

e nucleotide sequence of a primer used in the PCR amplification
r DNA fragment used in a method of the invention. The template
3 and the targeting vector was pSV-paz11. Host cells and/or
pressing Escherichia coli recE and recT genes (or functionally
nes) are useful in a novel cloning method, designated ET
hich involves homologous recombination via the recET mechanism
AAX87427-29). The method covalently combines one preferably
trachromosomal DNA fragment (the fragment to be cloned), with a
ferably circular DNA vector. The method is particularly
to complex or difficult DNA constructs, such as those intended
gous recombination in eukaryotic cells

7 BP; 19 A; 15 C; 21 G; 12 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 2; Length 67;
milarity 96.7%; Pred. No. 0.53;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||||
3AGCATCATCCTCTGCATGGTCAGGTCAT 32

tandard; DNA; 67 BP.

) (first entry)

P5603 PCR primer c-left.

, recE; recT; Escherichia coli; PCR; primer; pJP5603;
recombination; ss.

A2.

).

3; 98WO-EP007945.

7; 97EP-00121462.
3; 98EP-00118756.

3L EURO LAB MOLEKULARBIOLOGIE.

Zhang Y, Buchholz F;

457893/38.

or cloning DNA using a recE and recT homologous recombinant

; Page 137; 139pp; English.

e nucleotide sequence of a primer used in the PCR amplification
r DNA fragment used in a method of the invention. The template
3 and the targeting vector was pSV-paz11. Host cells and/or
pressing Escherichia coli recE and recT genes (or functionally
nes) are useful in a novel cloning method, designated ET
hich involves homologous recombination via the recET mechanism
AAX87427-29). The method covalently combines one preferably
trachromosomal DNA fragment (the fragment to be cloned), with a
ferably circular DNA vector. The method is particularly

CC applicable to complex or difficult DNA constructs, such as thos
CC for homologous recombination in eukaryotic cells
XX
SQ Sequence 67 BP; 19 A; 16 C; 18 G; 14 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 2; Length 67;
Best Local Similarity 96.7%; Pred. No. 0.53;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
Db 11 CGAGCATCATCCTCTGCATGGTCAGGTCAT 40

RESULT 28
ADC60826
ID ADC60826 standard; DNA; 100 BP.
XX
AC ADC60826;
XX
DT 18-DEC-2003 (first entry)
XX
DE Beta-galactosidase reporter sequence #11.
XX
KW Activated transcription factor; cis element; reporter sequence
KW beta-galactosidase; gene; ds.
XX
OS Escherichia coli.
XX
PN US2003143547-A1.
XX
PD 31-JUL-2003.
XX
PF 24-JAN-2002; 2002US-00057810.
XX
PR 24-JAN-2002; 2002US-00057810.
XX
PA (LIXX/) LI X.
PA (JIAN/) JIANG X.
XX
PI Li X, Jiang X;
XX
DR WPI; 2003-744884/70.
XX
PT Identifying multiple different activated transcription factors
PT sample by transducing or transfecting a cell sample and determ
PT activated transcription factors are present in the cell sample
XX
PS Disclosure; SEQ ID NO 41; 34pp; English.
XX
CC The invention relates to a method for identifying multiple dif
CC activated transcription factors in a cell sample, comprising t
CC or transfecting a cell sample to comprise a library of constru
CC forming mRNA transcription products by those of the transduced
CC transfected cells in which an activated transcription factor i
CC that binds to the cis element of the construct present in the
CC activates transcription of the reporter sequence of the constr
CC in the cell, determining which reporter sequences are comprise
CC the mRNA transcription products and determining which activate
CC transcription factors are present in the cell sample, based on
CC reporter sequences were transcribed. Each construct comprises
CC element sequence comprising one or more copies of a cis elemen
CC a transcription factor is capable of binding, the cis element
CC varying within the library of constructs, a promoter sequence
CC to the cis element sequence and a reporter sequence 3' relativ
CC promoter sequence that varies within the library, where the sa
CC element sequence is employed with a given reporter sequence wi
CC library of constructs. The method is useful for identifying mu
CC different activated transcription factors in a cell sample for
CC a disease state. This sequence represents a reporter sequence
CC method of the invention.
XX
SQ Sequence 100 BP; 22 A; 24 C; 32 G; 22 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 9; Length 100;
ilarity 96.7%; Pred. No. 0.61;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
AGCATCATCCTCTGCATGGTCAGGTCAT 76

andard; DB 1542 BP.

(first entry)

ssociated DNA fragment #2.

mmunodominant; glycoprotein D; antigen; serological testing;

lex virus.

2.

1.

); 2000RU-00101648.

); 2000RU-00101648.

TOR VIROLOGY & BIOTECHN RES CENTRE.

MA, Susloparov IM, Plyasunov IV;

94324/25.
49704.

plasmid DNA phsvd1 determining expression of gene US6
herpes simplex type-1 virus encoding immuno-dominant group of
n d(gd)hsv-1 in bacterium Escherichia coli cells.

Col 13-16; 9pp; Russian.

ion describes a novel recombinant plasmid DNA constructed in
ontaining the herpes simplex type-1 virus (HSV-1) US6 gene
ich encodes an immunodominant group of glycoprotein D (gD).
ptide shows the antigenic properties of herpes simplex virus
ified recombinant protein can be used as HSV-1 antigen for the
testing of HSV-1 in clinical practice. This sequence encodes
mplex type-1 virus US6-associated protein useful to the

42 BP; 325 A; 461 C; 431 G; 325 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 6; Length 1542;
ilarity 96.7%; Pred. No. 1.5;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
AGCATCATCCTCTGCATGGTCAGGTCAT 1103

andard; DNA; 3144 BP.

(first entry)

XX
DE beta-galactosidase reporter DNA fragment.
XX
KW Alpha-SMA; smooth muscle actin; screening; smooth muscle cell;
KW myofibroblast; gene therapy; beta-galactosidase; reporter; ds.
XX
OS Unidentified.
XX
PN EP1172375-A1.
XX
PD 16-JAN-2002.
XX
PF 22-DEC-2000; 2000EP-00128446.
XX
PR 11-JUL-2000; 2000DE-01033633.
PR 31-OCT-2000; 2000DE-01053879.
XX
PA (ODEN/) ODENTHAL M.
XX
PI Odenthal M, Jung D;
XX
DR WPI; 2002-149590/20.
XX
PT New nucleic acid containing regulatory region of the smooth mus
PT gene, useful e.g. for manipulating gene expression in smooth mu
PT cells.
XX
PS Disclosure; Page 21-22; 44pp; German.
XX
CC This invention describes a novel nucleic acid (I) comprising:
CC least one functional region (Ia) from the regulatory region of
CC smooth muscle actin (SMA) gene and (ii) at least one additional
CC functional sequence (Ib) operably linked to (Ia). The products
CC invention can be used for preparing genetically modified eukary
CC or organisms, for isolation and screening of smooth muscle cell
CC myofibroblasts or related cells, and for manipulation of gene e
CC and/or cell function in smooth muscle cell or myofibroblasts,
CC particularly for gene therapy. Component (Ia) provides cell-typ
CC differentiation-specific expression or modulation of genes. Thi
CC represents a beta-galactosidase DNA fragment which can be used
CC reporter molecule under the control of the alpha-SMA described
CC invention
XX
SQ Sequence 3144 BP; 694 A; 857 C; 908 G; 685 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 6; Length 3144;
Best Local Similarity 96.7%; Pred. No. 1.9;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
Db 1143 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1172

RESULT 31
AAH45081
ID AAH45081 standard; DNA; 3150 BP.
XX
AC AAH45081;
XX
DT 04-SEP-2001 (first entry)
XX
DE LacZ gene.
XX
KW LacZ; immunosuppressive; epigenetic regulation motif; immune re
KW T-cell response; methylation activity;
KW methyl DNA binding protein identification; ds.
XX
OS Escherichia coli.
XX
PN WO200140478-A2.
XX
PD 07-JUN-2001.

'0; 2000WO-EP012793.

'9; 99CA-02291367.

'ST PASTEUR.
'NT NAT RECH SCI.

Henry I, Choulika A;

367812/38.

ed polynucleotide having reduced or increased content of
: control motifs for studying, increasing and/or reducing gene
ı, and improving DNA vaccination methods.

Fig 1; 75pp; English.

ıt invention relates to modified LacZ genes (see AAH45079 and
in which epigenetic regulation motifs have been mutated
o the wild-type sequence (the present sequence). The mutant
: are useful for inducing in a second host, a protective immune
against a gene product of a first host. The mutant Lac Z genes
seful for evaluating a promoter in biological systems, for
methylation activity in biologically systems and/or for
g unknown methyl DNA binding proteins. The mutant LacZ genes
seful for compensating a genetic defect, and for therapeutic
ns. The mutant LacZ genes are also useful for minimising a T-
nse against the T-cells or tissues treated with them

150 BP; 706 A; 853 C; 904 G; 687 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 4; Length 3150;
milarity 96.7%; Pred. No. 1.9;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
TAGCATCATCCTCTGCATGGTCAGGTCAT 52
 ||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1178
```

tandard; DNA; 3157 BP.

2 (first entry)

tosidase gene.

ssion modulation system; gene expression cassette; promoter;
ation domain; DNA-binding domain; response element; LBD; DBD;
clear receptor ligand binding domain; gene therapy; TD; EcR;

a coli.

2-A2.

2.

2; 2002WO-US005090.

1; 2001US-0269799P.
1; 2001US-0313925P.

HM & HAAS CO.

Kumar MB, Cress DE, Fujimoto TT;

682765/73.

PT Novel gene expression modulation system useful for modulating
PT of a gene of interest, comprising a Group H nuclear receptor l
PT binding domain comprising a substitution mutation.
XX
PS Example 1; Page 118-120; 148pp; English.
XX
CC The present invention relates to a new gene expression modulat
CC The invention comprises a gene expression cassette, capable of
CC expressed in a host cell, comprising a polynucleotide that enc
CC polypeptide comprising a transactivation domain, a DNA-bindinc
CC that recognises a response element associated with a gene whos
CC expression is to be modulated, and a Group H nuclear receptor
CC binding domain (LBD). The invention is useful for modulating t
CC expression of the genE in a host cell which involves introduci
CC host cell, and introducing into the host cell a ligand, where
CC be modulated is a component of a gene expression cassette comp
CC response element by the DBD (DNA-binding domain), a promoter t
CC activated by the TD (transactivation domain) and a gene whose
CC is to be modulated, where upon introduction of the ligand intc
CC cell, expression of the desired gene is modulated. The inducib
CC expression system and its use for modulating gene expression i
CC cell overcome the limitations of currently available inducible
CC systems and provides an effective means for controlling gene e
CC The present nucleic acid sequence represents a DNA sequence th
CC in the methods of the invention
XX
SQ Sequence 3157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 6; Length 3157;
Best Local Similarity 96.7%; Pred. No. 1.9;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

```
Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              | ||||||||||||||||||||||||||||
Db       1170 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1199
```

RESULT 33
ACC00558
ID ACC00558 standard; DNA; 3157 BP.
XX
AC ACC00558;
XX
DT 23-JUN-2003 (first entry)
XX
DE E. coli lacZ reporter gene.
XX
KW Ecdysone receptor; BaEcr; pesticide; insect; male sterility; w
KW lacZ; ds.
XX
OS Escherichia coli.
XX
PN WO2003027266-A2.
XX
PD 03-APR-2003.
XX
PF 20-FEB-2002; 2002WO-US005234.
XX
PR 26-SEP-2001; 2001US-0325534P.
XX
PA (ROHM) ROHM & HAAS CO.
XX
PI Zhang J, Cress DE, Palli SR, Dhadialla TS;
XX
DR WPI; 2003-342734/32.
XX
PT Novel isolated whitefly (Bamecia argentifoli) ecdysone recepto
PT polypeptide, useful for screening for molecules e.g., agonist
PT antagonist of ecdysone receptor activity in a cell, which are
PT pesticides.
XX
PS Example 2; Page 76-78; 85pp; English.
XX

ion relates to a whitefly (Bamecia argentifoli ecdysone
BaEcr)) polypeptide and encoding polynucleotides. The BaEcr
e and polynucleotides are useful for modulating gene expression
sone receptor-based gene expression system, and for identifying
ing ligands specific for binding to a ligand-binding domain of
for identifying and selecting compounds exhibiting specific
the ligand binding domain of BaEcr. Agonists and antagonists
odulate insect physiology and development, thus modifying the
poral sequence of developmental agents in an insect. For
ccelerating insect development can be achieved for instance in
sterile males. Alternatively, it may be useful slow development
festation, such that the insects reach destructive stages of
t only after commercial crops may have passed sensitive stages.
ts and antagonists are also useful for artificially maintaining
a specific developmental stage. The present sequence
an E. coli lacZ reporter gene, used in the construction of
aEcr gene expression cassettes

157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 7; Length 3157;
milarity 96.7%; Pred. No. 1.9;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1199

tandard; DNA; 3157 BP.

3 (first entry)

cZ reporter gene.

; leafhopper; ecdysone receptor; homopteran; insecticide;
ssion; lacZ; gene; ds.

a coli.

39-A1.

3.

2; 2002WO-US005026.

1; 2001US-0325096P.

HM & HAAS CO.

181913/45.

ated leafhopper ecdysone receptor polypeptide from homopteran
lopper e.g. Nephotetix cincticeps, for identifying molecules
ate receptor activity and in insecticide discovery.

Page 71-73; 81pp; English.

ion relates to an isolated leafhopper ecdysone receptor
e (NcEcR) from the homopteran green leafhopper (Nephotetix
and encoding polynucleotide. The NcEcR polypeptide,
tide and vectors comprising the polynucleotide are useful in
of gene expression and insecticide discovery. The NcEcR
e and polynucleotide are useful for regulating gene expression
eptide of interest in a host cell and in identifying new
hat modulate activity of a leafhopper ecdysone receptor (EcR).
tibodies are useful as a reagent for determining qualitative

CC or quantitative presence of the receptor and to separate or pu
CC receptor. A composition comprising NcEcR polypeptide, polynucl
CC vectors are useful for formulation of biological material for
CC gene expression modulation system or a ligand-screening assay.
CC present sequence represents an E. coli LacZ reporter gene frag
CC in the construction of leafhopper ecdysone receptor gene expre
CC cassettes
XX
SQ Sequence 3157 BP; 697 A; 856 C; 915 G; 689 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 8; Length 3157;
Best Local Similarity 96.7%; Pred. No. 1.9;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
Db 1170 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1199

RESULT 35
AAD22647
ID AAD22647 standard; DNA; 3279 BP.
XX
AC AAD22647;
XX
DT 26-FEB-2002 (first entry)
XX
DE Beta-galactosidase reporter gene.
XX
KW Beta-galactosidase; gene expression; human immunodeficiency vi
KW psi gene; nucleocapsid protein; HIV inhibitor screening; ds.
XX
OS Unidentified.
XX
PN WO200177312-A1.
XX
PD 18-OCT-2001.
XX
PF 18-OCT-2000; 2000WO-KR001173.
XX
PR 08-APR-2000; 2000KR-00018489.
XX
PA (YOOJ/) YOO J C.
XX
PI Nam H, Kim S;
XX
DR WPI; 2002-025900/03.
XX
PT Novel microorganism for screening HIV inhibitors, is cotransfo
PT plasmid vector (PV) containing gene expressing HIV nucleocapsi
PT and PV containing HIV psi gene and beta-galactosidase reporter
XX
PS Claim 7; Page 26-29; 35pp; English.
XX
CC The invention relates to a microorganism cotransformed with a
CC expressing human immunodeficiency virus (HIV) nucleocapsid pro
CC plasmid vector containing HIV psi gene and beta-galactosidase
CC gene. The microorganism is useful for screening HIV packaging
CC by culturing microorganism, treating microorganism with putati
CC compounds or compositions of HIV inhibitors, and measuring the
CC change in beta-galactosidase expression in the culture. Microo
CC useful for high throughput screening (HTS) of HIV inhibitors.
CC sequence is beta-galactosidase reporter gene which is used for
CC constructing pNH1 (SL1234) plasmid
XX
SQ Sequence 3279 BP; 730 A; 888 C; 940 G; 721 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 6; Length 3279;
Best Local Similarity 96.7%; Pred. No. 2;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||

CGAGCATCATCCTCTGCATGGTCAGGTCAT 1307

standard; DNA; 3284 BP.

04 (first entry)

a-2 adrenergic receptor construct.

catic; gene therapy; minicell; membrane protein; cancer; human.

ens.

099-A1.

03.

02; 2002US-00157305.

01; 2001US-0295566P.
02; 2002US-0359843P.

ABBADINI R A.
URBER M W.
ERKLEY N.
EGALL A M.
LEPPER R.

RA, Surber MW, Berkley N, Segall AM, Klepper R;

-711671/67.

ell comprising a membrane protein consisting of eukaryotic, terial protein or organellar protein, useful for preparing a on for treating or preventing e.g. cancer.

; Page 219-220; 0pp; English.

ion relates to a new minicell comprising a membrane protein eukaryotic membrane protein, archeabacterial membrane protein llar membrane protein. The minicell comprises membrane membrane fusion protein, eubacterial minicell, poroplast, t, protoplast, biologically active compound or expression where the first expression construct comprises expression operably linked to an ORF (open reading frame) that encodes a protein. It comprises a second expression construct, having operably linked to a gene. The expression sequences are and/or repressible. The membrane conjugate comprises a membrane emically linked to a conjugated compound. The conjugated comprises nucleic acid, polypeptide, lipid or small molecule. product of the gene is a nucleic acid or polypeptide and the expression of the ORF that encodes the protein. The le is a membrane protein, soluble protein or secreted protein. ne protein is a membrane fusion protein comprising a first le, comprising at least one transmembrane domain or at least one ane anchoring domain, and a second polypeptide. The second le is not derived from a eubacterial protein and is neither a r a glutathione-S-transferase polypeptide. The minicell is preparing a composition for treating or preventing cancer. The equence is a DNA expression construct used to test the minicell vention, comprising a human gene (or fragment).

284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 9; Length 3284;
milarity 96.7%; Pred. No. 2;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
              |||||||||||||||||||||||||||||
Db       1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283
```

RESULT 37
ADC24539
ID ADC24539 standard; cDNA; 3284 BP.
XX
AC ADC24539;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human beta 2-adrinergic receptor cDNA.
XX
KW ss; minicell; episomal expression construct; cancer; asthma;
KW inflammation; rheumatoid arthritis; diabetes; Alzheimer's dis
KW Parkinson's disease; HIV; bacterial infection; hepatitis;
KW myocardial ischaemia; human; gene.
XX
OS Homo sapiens.
XX
PN US2003190749-A1.
XX
PD 09-OCT-2003.
XX
PF 28-MAY-2002; 2002US-00157215.
XX
PR 24-MAY-2001; 2001US-0293566P.
PR 25-FEB-2002; 2002US-0359843P.
PR 24-MAY-2002; 2002US-00154951.
XX
PA (SURB/) SURBER M W.
PA (SABB/) SABBADINI R A.
PA (SEGA/) SEGALL A M.
PA (BERK/) BERKLEY N.
XX
PI Surber MW, Sabbadini RA, Segall AM, Berkley N;
XX
DR WPI; 2003-831632/77.
XX
PT New minicell-producing parent cell comprising an expression el
PT mutation in an endogenous gene, useful for producing achromos
PT anucleate cells for diagnostic or therapeutic purposes and fo
PT discovery.
XX
PS Example 22; SEQ ID NO 186; 242pp; English.
XX
CC The invention relates to a minicell-producing parent cell. The
CC cell comprises: an expression element that comprises a gene o
CC linked to expression sequences that are inducible and/or repre
CC where induction or repression of the gene regulates the copy r
CC episomal expression construct and/or causes or enhances the p
CC minicells; and/or a mutation in an endogenous gene, where the
CC regulates the copy number of an episomal expression construct
CC causes or enhances minicell production. Also disclosed are con
CC and methods for preparing the minicells (or a soluble and/or
CC protein, or antibodies and/or antibody derivatives that recogn
CC immunogenic epitope on the native form of a membrane protein,
CC associating a radioactive compound with a cell), a method of t
CC a membrane protein from a minicell membrane to a biological me
CC pharmaceutical composition comprising the minicell, a method c
CC the above pharmaceutical composition, a method of detecting an
CC is specifically bound by a binding moiety, a method of in situ
CC a tissue or organ, methods of determining the rate or amount c
CC of nucleic acid from a minicell to a cell, a method of determi
CC three-dimensional structure of a membrane protein, a method of
CC identifying ligand-interacting atoms in a defined three-dimens
CC structure of a target protein, methods of identifying a nuclei
CC encodes the above protein, and methods of bioremediation. The
CC producing parent cell is useful for producing achromosomal and
CC cells for diagnostic and therapeutic applications (e.g. in dia
CC treating cancer, asthma, allergies, inflammation, rheumatoid a

Alzheimer's disease, Parkinson's disease, HIV, bacterial
, hepatitis or myocardial ischaemia), as well as research tools
for drug discovery or for delivery of nucleic acids and other
compounds to cells. The present sequence is a human cDNA used
binant construct for a mini-cell of the invention. Note: The
ve mixed up the seq id numbers between the disclosure and the
isting. This means that several of the sequences cannot be
ly identified and some of the rest may be mis-identified.

284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 9; Length 3284;
milarity 96.7%; Pred. No. 2;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

tandard; DNA; 3284 BP.

4 (first entry)

-2 adrenergic receptor, SalI-XhoI fragment.

ds; membrane protein; transmembrane domain;
nchoring domain; Type III secretion system; achromosomal cell;
cell; cancer; asthma; allergy; inflammation;
arthritis; diabetes; Alzheimer's disease; Parkinson's disease;
ion; bacterial infection; hepatitis; myocardial ischaemia;

ns.

14-A1.

3.

2; 2002US-00157299.

1; 2001US-0295566P.
2; 2002US-0359843P.

BBADINI R A.
RKLEY N.
RBER M W.

RA, Berkley N, Surber MW;

844449/78.

ll useful for producing achromosomal and anucleate cells for
or treating e.g. cancer, asthma, allergies, inflammation,
Alzheimer's disease or HIV, and as research tools and agents
iscovery.

; SEQ ID NO 186; 244pp; English.

ion relates to a minicell comprising at least one nucleic acid.
ll displays a binding moiety directed to a target compound,
binding moiety is selected from a eukaryotic membrane protein,
acterial membrane protein, an organellar membrane protein, and
rotein. The fusion protein comprises a first polypeptide
at least one transmembrane domain or at least one membrane
domain, and a second polypeptide that is not derived from a
l protein and is neither a His tag nor a glutathione-S-
e polypeptide, where the polypeptide comprises a binding
so included is the method of introducing a nucleic acid into a

CC cell, comprising contacting the cell with the minicell cited a
CC minicell is selected from a eubacterial minicell, a poroplast,
CC spheroplast and a protoplast. The nucleic acid comprises an ex
CC construct comprising expression sequences operably linked to a
CC encoding the proteins mentioned above or encoding a therapeuti
CC polypeptide. The therapeutic polypeptide is a membrane polypep
CC soluble polypeptide. The soluble polypeptide comprises a cellu
CC secretion sequence. The expression sequences are inducible and
CC repressible. These are induced and/or depressed when the bindi
CC displayed by the minicell binds to its target compound. The OR
CC polypeptide having an amino acid sequence that facilitates cel
CC transfer of a biologically active compound contained within or
CC by the minicell. The membrane of the minicell comprises a syst
CC transferring a molecule from the interior of a minicell into t
CC cytoplasm of the cell. The system is a Type III secretion syst
CC minicell and method are useful in producing achromosomal and a
CC cells for diagnostic and therapeutic applications (e.g. in dia
CC treating cancer, asthma, allergies, inflammation, rheumatoid a
CC diabetes, Alzheimer's disease, Parkinson's disease, HIV, bacte
CC infections, hepatitis or myocardial ischaemia), as well as res
CC and agents for drug discovery or for delivery of nucleic acids
CC bioactive compounds to cells. The present sequence is a human
CC sequence from a gene of interest, incorporated into a minicell
CC of the invention.
XX
SQ Sequence 3284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 9; Length 3284;
Best Local Similarity 96.7%; Pred. No. 2;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

RESULT 39
ADE10522
ID ADE10522 standard; DNA; 3284 BP.
XX
AC ADE10522;
XX
DT 29-JAN-2004 (first entry)
XX
DE Minicell associated DNA #46.
XX
KW membrane protein transfer; minicell membrane; biological membr
KW hyperproliferative disorder; cancer; ds.
XX
OS Synthetic.
XX
PN US2003199089-A1.
XX
PD 23-OCT-2003.
XX
PF 28-MAY-2002; 2002US-00157318.
XX
PR 05-JUN-2001; 2001US-0295566P.
PR 25-FEB-2002; 2002US-0359843P.
XX
PA (SURB/) SURBER M W.
PA (SABB/) SABBADINI R A.
XX
PI Surber MW, Sabbadini RA;
XX
DR WPI; 2003-852795/79.
XX
PT Transferring a membrane protein from a minicell membrane to a
PT membrane for diagnosing or treating e.g. cancer by allowing th
PT and biological membrane to remain in contact for a sufficient
PT the transfer to occur.
XX
PS Disclosure; SEQ ID NO 186; 243pp; English.

ion relates to a method of transferring a membrane protein from
membrane to a biological membrane which comprises contacting a
o the biological membrane and allowing the minicell and
membrane to remain in contact for a period of time sufficient
ansfer to occur. The method is useful for transferring a
rotein from a minicell membrane to a biological membrane for
a composition for diagnosing or treating hyperproliferative
e.g. cancer. The present sequence represents a minicell
DNA.

284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

36.9%; Score 28.4; DB 9; Length 3284;
milarity 96.7%; Pred. No. 2;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 1283

tandard; DNA; 3284 BP.

4 (first entry)

-2 adrenergic receptor, SalI-XhoI fragment.

eukaryotic expression sequence; open reading frame; ORF;
l minicell; poroplast; spheroplast; protoplast;
al cell; anucleate cell; drug discovery; ds; human.

ns.

79-A1.

3.

2; 2002US-00157391.

1; 2001US-0293566P.
2; 2002US-0359843P.
2; 2002US-00154951.

BBADINI R A.
RKLEY N.

RA, Berkley N;

874920/81.

g the rate of transfer of nucleic acid from a minicell to a
vice versa) useful in the production of achromosomal and
cells used for diagnostic and therapeutic applications.

2; SEQ ID NO 186; 242pp; English.

ion relates to determining the rate of transfer of nucleic acid
nicell to a cell, determining the amount of a nucleic acid
ed to a cell from a minicell and detecting the expression of an
n element in a cell. The minicell comprises an expression
aving eukaryotic expression sequences operably linked to an open
rame (ORF) encoding a detectable polypeptide, the minicells
binding group and the binding group displays an epitope of the
minicell is a eubacterial minicell, a poroplast, a spheroplast
oplast. The cell is a eukaryotic cell. The binding group is an
or antibody derivative, especially a single-chain antibody, an
r an organic compound. The detectable polypeptide is a
nt polypeptide. The methods are used in the production of

CC achromosomal and anucleate cells useful for applications such
CC diagnostic and therapeutic uses, as well as research tools and
CC drug discovery. The present sequence is a human DNA sequence u
CC minicell construct of the invention. Note: The authors have mi
CC SEQ ID numbers between the text and the sequence listing such
CC of the sequences cannot be conclusively identified.
XX
SQ Sequence 3284 BP; 753 A; 865 C; 927 G; 739 T; 0 U; 0 Other;

Query Match 36.9%; Score 28.4; DB 9; Length 3284;
Best Local Similarity 96.7%; Pred. No. 2;
Matches 29; Conservative 0; Mismatches 1; Indels 0;

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
|||||||||||||||||||||||||||||
Db 1254 CGAGCATCATCCTCTGCATGGTCAGGTCAT 1283

Search completed: April 13, 2004, 18:34:55
Job time : 220.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 17:11:06 ; Search time 1107.5 Seconds
(without alignments)
3013.466 Million cell updates/sec

US-10-090-326-25
77
1 acttacncagcccagcanca.....ggagnacaacagccancagt 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

3470272 seqs, 21671516995 residues

hits satisfying chosen parameters: 6940544

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

). is the number of results predicted by chance to have a score greater than or equal to the score of the result being and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Descript
1	30	39.0	1059	6	AR061745	AR061745
2	30	39.0	1059	6	AR061746	AR061746
3	30	39.0	1059	6	AR061986	AR061986
4	30	39.0	1059	6	AR061987	AR061987
5	30	39.0	1059	6	AR206220	AR206220
6	30	39.0	1059	6	AR206221	AR206221
7	30	39.0	1452	9	BC027179	BC027179
8	30	39.0	1587	6	AR061747	AR061747
9	30	39.0	1587	6	AR061748	AR061748
10	30	39.0	1587	6	AR061988	AR061988
11	30	39.0	1587	6	AR061989	AR061989
12	30	39.0	1587	6	E59483	E59483 M
13	30	39.0	1587	6	AR206222	AR206222
14	30	39.0	1587	6	AR206223	AR206223
15	30	39.0	1590	6	AX113825	AX113825
16	30	39.0	1590	9	HSU01873	U01873 H
17	30	39.0	1880	9	HUMTYR	J03581 H
18	30	39.0	1886	6	I00224	I00224 S
19	30	39.0	1888	9	HSTYROR	Y00819 H
20	30	39.0	1894	6	AR004653	AR004653
21	30	39.0	1894	6	AR062126	AR062126
22	30	39.0	1894	6	AR062127	AR062127
23	30	39.0	1894	6	I21255	I21255 S
24	30	39.0	1906	6	AR003571	AR003571
25	30	39.0	1906	6	I17316	I17316 S
26	30	39.0	1910	6	AR061264	AR061264
27	30	39.0	1929	9	HUMTYRM	M74314 H
28	30	39.0	2041	6	AX285154	AX285154
29	30	39.0	2384	6	AR167367	AR167367
30	30	39.0	2384	6	E11466	E11466 C
31	30	39.0	2384	6	I49609	I49609 C
32	30	39.0	2384	9	HUMTYRA	M27160 H
C 33	28.8	37.4	155690	10	AC121885	AC1218
C 34	28.6	37.1	162383	2	AC118220	AC11822
35	28.4	36.9	67	6	BD131906	BD13190
36	28.4	36.9	67	6	BD131908	BD13190
37	28.4	36.9	299	4	AB128155	AB12815
38	28.4	36.9	1063	1	ECO487615	AJ48761
39	28.4	36.9	1063	1	ECO487616	AJ48761
40	28.4	36.9	1063	1	ECO487617	AJ48761
41	28.4	36.9	1950	6	I00682	I00682
42	28.4	36.9	3078	1	ECLACZ	V00296
43	28.4	36.9	3096	6	AR387286	AR38728
44	28.4	36.9	3144	6	AX348046	AX34804
45	28.4	36.9	3144	6	AX353913	AX35391

ALIGNMENTS

RESULT 1
AR061745
LOCUS AR061745 1059 bp DNA linear PA
DEFINITION Sequence 58 from patent US 5843654.
ACCESSION AR061745
VERSION AR061745.1 GI:5989436
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Heisler,L.M., Fors,L. and Brow,M.Ann.D.
TITLE Rapid detection of mutations in the p53 gene
JOURNAL Patent: US 5843654-A 58 01-DEC-1998;
FEATURES Location/Qualifiers

1. .1059
/organism="unknown"
/mol_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;
imilarity 67.9%; Pred. No. 1.9;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| |||||| | | || | | | |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 552

)61746 1059 bp DNA linear PAT 29-SEP-1999
quence 59 from patent US 5843654.
)61746
'61746.1 GI:5989437

nown.
nown.
lassified.
(bases 1 to 1059)
sler,L.M., Fors,L. and Brow,M.Ann.D.
id detection of mutations in the p53 gene
ent: US 5843654-A 59 01-DEC-1998;
Location/Qualifiers
1. .1059
/organism="unknown"
/mol_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;
milarity 67.9%; Pred. No. 1.9;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| |||||| | | || | | | |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 552

;1986 1059 bp DNA linear PAT 29-SEP-1999
ience 58 from patent US 5843669.
;1986
;1986.1 GI:5989677

own.
own.
assified.
bases 1 to 1059)
ser,M.W., Lyamichev,V.I. and Lyamichev,N.
vage of nucleic acid acid using thermostable methoanococcus
aschii FEN-1 endonucleases
nt: US 5843669-A 58 01-DEC-1998;
Location/Qualifiers
1. .1059
/organism="unknown"
/mol_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1059;

Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | || | | | |||||||
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 535 AGTACAACAGCCATCAGT 552

RESULT 4
AR061987
LOCUS AR061987 1059 bp DNA linear PA'
DEFINITION Sequence 59 from patent US 5843669.
ACCESSION AR061987
VERSION AR061987.1 GI:5989678
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Kaiser,M.W., Lyamichev,V.I. and Lyamichev,N.
TITLE Cleavage of nucleic acid acid using thermostable methoa
jannaschii FEN-1 endonucleases
JOURNAL Patent: US 5843669-A 59 01-DEC-1998;
FEATURES Location/Qualifiers
source 1. .1059
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 1059;
Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
|||||| |||||||||| |||||| | | || | | | |||||||
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 535 AGTACAACAGCCATCAGT 552

RESULT 5
AR206220
LOCUS AR206220 1059 bp DNA linear PAT
DEFINITION Sequence 58 from patent US 6372424.
ACCESSION AR206220
VERSION AR206220.1 GI:21504758
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Brow,M.Ann.D., Lyamichev,V.I. and Olive,D.Michael.
TITLE Rapid detection and identification of pathogens
JOURNAL Patent: US 6372424-A 58 16-APR-2002;
FEATURES Location/Qualifiers
source 1. .1059
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 1059;
Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATT
|||||| |||||||||| |||||| | | || | | | |||||||
Db 475 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGAT

```
GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 552
```

```
06221                 1059 bp    DNA     linear   PAT 20-JUN-2002
uence 59 from patent US 6372424.
06221
06221.1  GI:21504759

nown.
nown.
lassified.
(bases 1 to 1059)
w,M.Ann.D., Lyamichev,V.I. and Olive,D.Michael.
id detection and identification of pathogens
ent: US 6372424-A 59 16-APR-2002;
      Location/Qualifiers
      1. .1059
      /organism="unknown"
      /mol_type="unassigned DNA"


          39.0%;  Score 30;  DB 6;  Length 1059;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 534

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 552
```

```
27179                 1452 bp    mRNA    linear   PRI 06-OCT-2003
o sapiens tyrosinase (oculocutaneous albinism IA), mRNA (cDNA
ne MGC:9191 IMAGE:3923096), complete cds.
27179
27179.1  GI:20072605
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 1452)
ausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
usner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
schul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
kins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
tchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
pleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
eetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
ninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
amson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
ernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
ley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
lalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
ey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
chez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
ffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
kson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
terfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
nerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
eration and initial analysis of more than 15,000 full-length
an and mouse cDNA sequences
c. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
```

```
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1452)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2002) National Institutes of Health,
            Gene Collection (MGC), Cancer Genomics Office, National
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 2(
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN
            DNA Sequencing by: Baylor College of Medicine Human Ger
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lou
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Ne
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information
            through the I.M.A.G.E. Consortium/LLNL at: http://image
            Series: IRAK Plate: 26 Row: d Column: 20
            This clone was selected for full length sequencing beca
            passed the following selection criteria: matched mRNA g
FEATURES              Location/Qualifiers
     source           1. .1452
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:9191 IMAGE:3923096"
                      /tissue_type="Skin, melanotic melanoma."
                      /clone_lib="NIH_MGC_72"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     gene             1. .1452
                      /gene="TYR"
                      /note="synonyms: OCAIA, OCA1A"
                      /db_xref="LocusID:7299"
                      /db_xref="MIM:606933"
     CDS              62. .1195
                      /codon_start=1
                      /product="TYR protein"
                      /protein_id="AAH27179.1"
                      /db_xref="GI:20072606"
                      /db_xref="LocusID:7299"
                      /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLME
                      PCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTCQ
                      CKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSD
                      NGSTPMFNDINIYDLFVWMHYYVSMDALLGGSEIWRDIDFAHEAPA
                      EQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNLLS
                      SRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPRLPSSADVEFCL
                      AANFSFRNTLEEMGFLHVGWAGLKLLTSRDPPPWPPKMLGLQA"
     misc_feature     392. .1096
                      /note="tyrosinase; Region: Common central doma
                      tyrosinase. This family also contains polyphen
                      and some hemocyanins. Binds two copper ions vi
                      of three histidines. This family is related to
                      /db_xref="CDD:pfam00264"
ORIGIN

  Query Match             39.0%;  Score 30;  DB 9;  Length 1452;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        843 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA
```

```
GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 920
```

```
61747                 1587 bp    DNA     linear   PAT 29-SEP-1999
uence 60 from patent US 5843654.
61747
61747.1  GI:5989438

nown.
nown.
lassified.
 (bases 1 to 1587)
sler,L.M., Fors,L. and Brow,M.Ann.D.
id detection of mutations in the p53 gene
ent: US 5843654-A 60 01-DEC-1998;
       Location/Qualifiers
       1. .1587
       /organism="unknown"
       /mol_type="unassigned DNA"


           39.0%;  Score 30;  DB 6;  Length 1587;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| |||||||||| ||||
GTACAACAGCCATCAGT 859
```

```
061748                1587 bp    DNA     linear   PAT 29-SEP-1999
quence 61 from patent US 5843654.
061748
061748.1  GI:5989439

nown.
nown.
lassified.
 (bases 1 to 1587)
sler,L.M., Fors,L. and Brow,M.Ann.D.
id detection of mutations in the p53 gene
ent: US 5843654-A 61 01-DEC-1998;
       Location/Qualifiers
       1. .1587
       /organism="unknown"
       /mol_type="unassigned DNA"


           39.0%;  Score 30;  DB 6;  Length 1587;
imilarity   67.9%;  Pred. No. 1.9;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 859
```

```
LOCUS       AR061988                1587 bp    DNA     linear   PAT
DEFINITION  Sequence 60 from patent US 5843669.
ACCESSION   AR061988
VERSION     AR061988.1  GI:5989679
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1587)
  AUTHORS   Kaiser,M.W., Lyamichev,V.I. and Lyamichev,N.
  TITLE     Cleavage of nucleic acid acid using thermostable metho
            jannaschii FEN-1 endonucleases
  JOURNAL   Patent: US 5843669-A 60 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .1587
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             39.0%;  Score 30;  DB 6;  Length 1587;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 11
AR061989
LOCUS       AR061989                1587 bp    DNA     linear   PAT
DEFINITION  Sequence 61 from patent US 5843669.
ACCESSION   AR061989
VERSION     AR061989.1  GI:5989680
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1587)
  AUTHORS   Kaiser,M.W., Lyamichev,V.I. and Lyamichev,N.
  TITLE     Cleavage of nucleic acid acid using thermostable metho
            jannaschii FEN-1 endonucleases
  JOURNAL   Patent: US 5843669-A 61 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .1587
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             39.0%;  Score 30;  DB 6;  Length 1587;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 12
E59483
LOCUS       E59483                  1587 bp    DNA     linear   PA
DEFINITION  Method for determing effect of gray hair-preventing ag
ACCESSION   E59483
VERSION     E59483.1  GI:18622554
KEYWORDS    JP 2000300298-A/3.
```

```
no sapiens (human)
no sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 1587)
eta,T. and Ifuku,O.
thod for determing effect of gray hair-preventing agent
ent: JP 2000300298-A 3 31-OCT-2000;
SEIDO CO LTD
   Homo sapiens (human)
   JP 2000300298-A/3
   31-OCT-2000
   23-APR-1999 JP 1999117006

   TATSURO IDETA,OJI IFUKU
   C12Q1/68,C12N15/09,G01N33/15,G01N33/50,C12N15/00 CC
   Key             Location/Qualifiers
   source          1. .1587
                   /organism='Homo sapiens (human)'.
       Location/Qualifiers
       1. .1587
       /organism="Homo sapiens"
       /mol_type="genomic DNA"
       /db_xref="taxon:9606"


           39.0%;  Score 30;  DB 6;  Length 1587;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 859



06222                1587 bp    DNA     linear   PAT 20-JUN-2002
uence 60 from patent US 6372424.
06222
06222.1  GI:21504760

nown.
nown.
lassified.
(bases 1 to 1587)
w,M.Ann.D., Lyamichev,V.I. and Olive,D.Michael.
id detection and identification of pathogens
ent: US 6372424-A 60 16-APR-2002;
      Location/Qualifiers
      1. .1587
      /organism="unknown"
      /mol_type="unassigned DNA"

.
           39.0%;  Score 30;  DB 6;  Length 1587;
nilarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

GNACAACAGCCANCAGT 77
| ||||||||||| ||||
GTACAACAGCCATCAGT 859
```

```
AR206223
LOCUS       AR206223               1587 bp    DNA     linear   PA
DEFINITION  Sequence 61 from patent US 6372424.
ACCESSION   AR206223
VERSION     AR206223.1  GI:21504761
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1587)
  AUTHORS   Brow,M.Ann.D., Lyamichev,V.I. and Olive,D.Michael.
  TITLE     Rapid detection and identification of pathogens
  JOURNAL   Patent: US 6372424-A 61 16-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1587
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           39.0%;  Score 30;  DB 6;  Length 1587;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| |||||       |  | |  ||  |   | ||
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 15
AX113825
LOCUS       AX113825               1590 bp    DNA     linear   PAT
DEFINITION  Sequence 1 from Patent WO0127295.
ACCESSION   AX113825
VERSION     AX113825.1  GI:13939991
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hc
REFERENCE   1
  AUTHORS   Schadendorf,D., Paschen,A., Chakraborty,T. and Domann,E
  TITLE     Recombinant attenuated listerias for immunotherapy
  JOURNAL   Patent: WO 0127295-A 1 19-APR-2001;
            Deutsches Krebsforschungszentrum Stiftung des Oeffentli
            (DE)
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1590
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37863.1"
                     /db_xref="GI:13939992"
                     /db_xref="REMTREMBL:CAC37863"
                     /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLME
                     PCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTCQ
                     CKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSD
                     NGSTPMFNDINIYDLFVWMHYYVSMDALLGGYEIWRDIDFAHEAPA
                     EQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNLLS
                     SRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPRLPSSADVEFCL
                     AANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQVQGS
                     VDSIFEQWLRRHRPLQEVYPEANAPIGHNRESYMVPFIPLYRNGDF
                     YLQDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALLAGLV
                     EEKQPLLMEKEDYHSLYQSHL"
ORIGIN

  Query Match           39.0%;  Score 30;  DB 6;  Length 1590;
```

```
imilarity   67.9%;  Pred. No. 1.9;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |  | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 841

AGNACAACAGCCANCAGT 77
|| ||||||||||| ||||
AGTACAACAGCCATCAGT 859
```

```
U01873                 1590 bp    mRNA    linear   PRI 27-SEP-1993
man tyrosinase mRNA, complete cds.
1873
1873.1  GI:403421

mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 1590)
ichard,V., Van Pel,A., Wolfel,T., Wolfel,C., De Plaen,E.,
the,B., Coulie,P. and Boon,T.
e tyrosinase gene codes for an antigen recognized by autologous
tolytic T lymphocytes on HLA-A2 melanomas
 Exp. Med. 178 (2), 489-495 (1993)
340625
40755
 (bases 1 to 1590)
ichard,B., Fuller,B.B., Vijayasaradhi,S. and Houghton,A.N.
duction of pigmentation in mouse fibroblasts by expression of
man tyrosinase cDNA
 Exp. Med. 169 (6), 2029-2042 (1989)
279151
99655
 (bases 1 to 1590)
on,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.
olation and sequence of a cDNA clone for human tyrosinase that
os at the mouse c-albino locus
oc. Natl. Acad. Sci. U.S.A. 84 (21), 7473-7477 (1987)
041128
33263
 (bases 1 to 1590)
ichard,V.G.
rect Submission
omitted (16-SEP-1993) Vincent G Brichard, Ludwig Institute for
ncer Research, Avenue Hippocrate, 74, Brussels B-1200 Belgium
        Location/Qualifiers
        1. .1590
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        1. .1590
        /codon_start=1
        /product="tyrosinase"
        /protein_id="AAB60319.1"
        /db_xref="GI:403422"
        /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLMEKECCPPWSGDRS
        PCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTCQCSGNFMGFNCGN
        CKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSDYVIPIGTYGQMK
        NGSTPMFNDINIYDLFVWMHYYVSMDALLGGSEIWRDIDFAHEAPAFLPWHRLFLLRW
        EQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNLLSPASFFSSWQIVC
        SRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPRLPSSADVEFCLSLTQYESGSMDK
        AANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQVQGSANDPIFLLHHAF
        VDSIFEQWLQRHRPLQEVYPEANAPIGHNRESYMVPFIPLYRNGDFFISSKDLGYDYS
        YLQDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALLAGLVSLLCRHKRKQLP
        EEKQPLLMEKEDYHSLYQSHL"


             39.0%;  Score 30;  DB 9;  Length 1590;
```

```
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |  | |  ||  |   | |
Db        782 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || ||||||||||| ||||
Db        842 AGTACAACAGCCATCAGT 859


RESULT 17
HUMTYR
LOCUS       HUMTYR                  1880 bp    mRNA    linear   PR
DEFINITION  Human tyrosinase mRNA, 3' end.
ACCESSION   J03581
VERSION     J03581.1  GI:340027
KEYWORDS    tyrosinase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Kwon,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.
  TITLE     Isolation and sequence of a cDNA clone for human tyros
            maps at the mouse c-albino locus
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 84 (21), 7473-7477 (1987
  MEDLINE   88041128
   PUBMED   2823263
REFERENCE   2  (bases 1144 to 1190)
  AUTHORS   Kwon,B.S., Haq,A.K., Pomerantz,S.H. and Halaban,R.
  TITLE     Correction: Isolation and sequence of a cDNA clone for
            tyrosinase that maps at the mouse c-albino locus
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 85, 6352-6352 (1988)
COMMENT     Original source text: Human melanocyte, cDNA to mRNA,
            [2]  revises [1].
            Draft entry and printed copy of sequence for [1] kindl
            B.S.Kwon, 29-SEP-1987.
FEATURES             Location/Qualifiers
     source          1. .1880
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="11q14-q21"
     gene            1. .1880
                     /gene="TYR"
     mRNA            <1. .1880
                     /gene="TYR"
                     /product="tyrosinase mRNA"
     CDS             <1. .1683
                     /gene="TYR"
                     /note="tyrosinase precursor (EC 1.14.18.1)"
                     /codon_start=1
                     /protein_id="AAA61241.1"
                     /db_xref="GI:340028"
                     /db_xref="GDB:G00-120-476"
                     /translation="LLLAVLYCLLWSFQTSAGHFPRACVSSKNLMI
                     CGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTCQ(
                     KFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSDY
                     GSTPMFNDINIYDLFVWMHYYVSMDALLGGYEIWRDIDFAHEAPAI
                     QEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNLLSI
                     RLEEYNSHQSLCNGTPEGPLRRNPGNHDKSTTPRLPSSADVEFCLS
                     ANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGHVPGTGSAI
                     SIFEQWLQRHRPLQEVYPEANAPIGHNRESYMVPFIPLYRNGDFF]
                     QDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALLAGPVSI
                     KQPLLMEKEGLPQLVSEPFIKGLGNRVGPKSPDLTLTQSNVQVPEI
     sig_peptide     <1. .36
                     /gene="TYR"
                     /note="tyrosinase signal peptide"
     mat_peptide     37. .1680
                     /gene="TYR"
                     /product="tyrosinase"
```

bp upstream of BalI site; chromosome 11.

39.0%; Score 30; DB 9; Length 1880;
imilarity 67.9%; Pred. No. 1.9;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| |||||| | | || | | | |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 838

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 856

)224 1886 bp ss-DNA linear PAT 21-MAY-1993
quence 1 from Patent US 4898814.
)224
)224.1 GI:313968

known.
known.
classified.
(bases 1 to 1886)
on,B.S.
cDNA clone for human tyrosinase
tent: US 4898814-A 1 06-FEB-1990;
ald Guthrie Foundation for Medical Research, Inc.; Sayre, PA
Jul 30, 1993 this sequence version replaced gi:285511.
Location/Qualifiers
1. .1886
/organism="unknown"
/mol_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 1886;
milarity 67.9%; Pred. No. 1.9;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| |||||| | | || | | | |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 844

AGNACAACAGCCANCAGT 77
| |||||||||| ||||
AGTACAACAGCCATCAGT 862

YROR 1888 bp mRNA linear PRI 28-JUL-1995
an mRNA for tyrosinase (EC 1.14.18.1).
)819
)819.1 GI:37508
opheno1 monooxygenase; tyrosinase.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 1888)
chard,B., Fuller,B.B., Vijayasaradhi,S. and Houghton,A.N.
luction of pigmentation in mouse fibroblasts by expression of
an tyrosinase cDNA
Exp. Med. 169 (6), 2029-2042 (1989)
:79151
)9655
(bases 1 to 1888)
chard,B.
ect Submission
mitted (12-APR-1989) Bouchard B., Memorial Sloan Kettering
cer Center, 1250 First Avenue, New York, N.Y. 10028, USA

COMMENT The sequence overlaps with that reported by Kwon et. a
Natl. Acad. Sci. USA 84:7473-7477(1987) <J03581>.
FEATURES Location/Qualifiers
source 1. .1888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BBTY.1."
/tissue_type="human melanoma"
/clone_lib="lambda gt10"
mRNA <1. .1888
/product="tyrosinase mRNA"
CDS <1. .1596
/codon_start=1
/product="put. tyrosinase preprotein"
/protein_id="CAA68756.1"
/db_xref="GI:37509"
/db_xref="GOA:P14679"
/db_xref="SWISS-PROT:P14679"
/translation="GRMLLAVLYCLLWSFQTSAGHFPRACVSSKN
RSPCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNR
GNCKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTI
MKNGSTPMFNDINIYDLFVWIHYYVSMDALLGGYEIWRDIDFAHE
RWEQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPN
VCSRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPRLPSSADVE
DKAANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQV
AFVDSIFEQWLRRHRPLQEVYPEANAPIGHNRESYMVPFIPLYRN
YSYLQDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALLA
LPEEKQPLLMEKEDYHSLYQSHL"
sig_peptide <1. .57
/note="put. signal peptide (AA -19 to -1)"
mat_peptide 58. .1593
/product="tyrosinase (AA 1-512)"
ORIGIN

Query Match 39.0%; Score 30; DB 9; Length 1888;
Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
|||||| |||||||||| |||||| | | || | | |
Db 788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
Db 848 AGTACAACAGCCATCAGT 865

RESULT 20
AR004653
LOCUS AR004653 1894 bp DNA linear PA
DEFINITION Sequence 1 from patent US 5747271.
ACCESSION AR004653
VERSION AR004653.1 GI:3965532
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1894)
AUTHORS Boon-Falleur,T., Brichard,V., Van Pel,A., De Plaen,E.,
Renauld,J.-C. and Wolfel,T.
TITLE Method for identifying individuals suffering from a cel
abnormality some of whose abnormal cells present comple
HLA-A2/tyrosinase derived peptides, and methods for tre
individuals
JOURNAL Patent: US 5747271-A 1 05-MAY-1998;
FEATURES Location/Qualifiers
source 1. .1894
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 1894;

```
imilarity   67.9%;  Pred. No. 1.9;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 865



062126               1894 bp    DNA     linear   PAT 29-SEP-1999
quence 1 from patent US 5843687.
062126
062126.1  GI:5989817

known.
known.
classified.
 (bases 1 to 1894)
lfel,T., Pel,A.Van., Brichard,V. and Boon-Falleur,T.
olated, tyrosinase derived peptides and uses thereof
tent: US 5843687-A 1 01-DEC-1998;
       Location/Qualifiers
       1. .1894
       /organism="unknown"
       /mol_type="unassigned DNA"


           39.0%;  Score 30;  DB 6;  Length 1894;
imilarity   67.9%;  Pred. No. 1.9;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

AGNACAACAGCCANCAGT 77
|| |||||||||| ||||
AGTACAACAGCCATCAGT 865



062127               1894 bp    DNA     linear   PAT 29-SEP-1999
quence 1 from patent US 5843688.
062127
062127.1  GI:5989818

known.
known.
classified.
 (bases 1 to 1894)
lfel,T., Van Pel,A., Brichard,V., Boon-Falleur,T., DePlaen,E.,
ilie,P., Renauld,J.-C. and Lethe,B.
olated tyrosinase derived peptides and uses thereof
tent: US 5843688-A 1 01-DEC-1998;
       Location/Qualifiers
       1. .1894
       /organism="unknown"
       /mol_type="unassigned DNA"


           39.0%;  Score 30;  DB 6;  Length 1894;
imilarity   67.9%;  Pred. No. 1.9;
;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
|||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 847

Qy          60 AGNACAACAGCCANCAGT 77
               || |||||||||| ||||
Db         848 AGTACAACAGCCATCAGT 865


RESULT 23
I21255
LOCUS       I21255                 1894 bp    DNA     linear   PA
DEFINITION  Sequence 1 from patent US 5519117.
ACCESSION   I21255
VERSION     I21255.1  GI:1601609
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1894)
  AUTHORS   Wolfel,T., Pel,A.V., Brichard,V. and Boon-Falleur,T.
  TITLE     Isolated, tyrosinase derived peptides and uses thereof
  JOURNAL   Patent: US 5519117-A 1 21-MAY-1996;
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           39.0%;  Score 30;  DB 6;  Length 1894;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        788 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        848 AGTACAACAGCCATCAGT 865


RESULT 24
AR003571
LOCUS       AR003571               1906 bp    DNA     linear   PA
DEFINITION  Sequence 1 from patent US 5744316.
ACCESSION   AR003571
VERSION     AR003571.1  GI:3964830
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1906)
  AUTHORS   Lethe,B., Brichard,V., Van Pel,A., Wolfel,T. and Boon-
  TITLE     Isolated, tyrosinase derived peptides and uses thereof
  JOURNAL   Patent: US 5744316-A 1 28-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .1906
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           39.0%;  Score 30;  DB 6;  Length 1906;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        806 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db        866 AGTACAACAGCCATCAGT 883


RESULT 25
```

```
7316                    1906 bp    DNA     linear   PAT 03-APR-1996
quence 1 from patent US 5487974.
7316
7316.1  GI:1252224

known.
known.
classified.
 (bases 1 to 1906)
on-Falleur,T., Brichard,V., Van Pel,A., De Plaen,E., Coulie,P.,
auld,J.-C., Wolfel,T. and Lethe,B.
thod for detecting complexes containing human leukocyte antigen
 (HLA-A2) molecules and a tyrosinase drived peptide on abnormal
lls
ent: US 5487974-A 1 30-JAN-1996;
       Location/Qualifiers
       1. .1906
       /organism="unknown"
       /mol_type="unassigned DNA"


            39.0%;  Score 30;  DB 6;  Length 1906;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 865

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 883



61264                   1910 bp    DNA     linear   PAT 29-SEP-1999
uence 18 from patent US 5843648.
61264
61264.1  GI:5988955

nown.
nown.
lassified.
 (bases 1 to 1910)
bins,P.F. and Rosenberg,S.A.
 and tyrosinase melanoma antigens and their use in diagnostic
 therapeutic methods
ent: US 5843648-A 18 01-DEC-1998;
       Location/Qualifiers
       1. .1910
       /organism="unknown"
       /mol_type="unassigned DNA"


            39.0%;  Score 30;  DB 6;  Length 1910;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
||||| |||||||||| ||||||       |   | |  ||  |   | |  |||||||
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 869

GNACAACAGCCANCAGT 77
| ||||||||| ||||
GTACAACAGCCATCAGT 887



TYRM                    1929 bp    mRNA    linear   PRI 14-JAN-1995
apien tyrosinase and mutant tyrosinase, complete cds.

ACCESSION   M74314
VERSION     M74314.1  GI:340039
KEYWORDS    tyrosinase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   Chintamaneni,C.D., Halaban,R., Kobayashi,Y., Witkop,C.
            Kwon,B.S.
  TITLE     A single base insertion in the putative transmembrane
            tyrosinase gene as a cause for tyrosinase-negative ocu
            albinism
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 88 (12), 5272-5276 (1991
  MEDLINE   91271371
   PUBMED   1711223
COMMENT     Original source text: Homo sapiens cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .1929
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="11q14-q21"
     gene            1. .1929
                     /gene="TYR"
     mRNA            1. .1929
                     /gene="TYR"
     CDS             32. .1636
                     /gene="TYR"
                     /codon_start=1
                     /product="tyrosinase"
                     /protein_id="AAA61244.1"
                     /db_xref="GI:340040"
                     /db_xref="GDB:G00-120-476"
                     /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLMI
                     PCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTC(
                     CKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSI
                     NGSTPMFNDINIYDLFVWMHIYYVSMDALLGGYEIWRDIDFSAHEI
                     RWEQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNI
                     VCSRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSTTPRLPSSADVEI
                     MDKAANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQV
                     HAFVDSIFEQWLQRHRPLQEVYPEANAPIGHRNRESYMVPFIPLYI
                     YDYSYLQDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALI
                     RKQLPEEKQPLLMEKEDYHSLYQSHL"
     variation       1511. .1512
                     /gene="TYR"
                     /note="G00-120-476"
ORIGIN

  Query Match           39.0%;  Score 30;  DB 9;  Length 1929;
  Best Local Similarity   67.9%;  Pred. No. 1.9;
  Matches   53;  Conservative    0;  Mismatches   24;  Indels    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCAT
              |||||| |||||||||| ||||||       |   | |  ||  |   | |
Db        819 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGA

Qy         60 AGNACAACAGCCANCAGT 77
              || ||||||||| ||||
Db        879 AGTACAACAGCCATCAGT 896


RESULT 28
AX285154
LOCUS       AX285154                2041 bp    DNA     linear   PAT
DEFINITION  Sequence 959 from Patent WO0179556.
ACCESSION   AX285154
VERSION     AX285154.1  GI:17045842
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
```

llie,J., Brown,J.L., Bolt,A. and van Huffel,C.
vel genes, compositions and methods for the identification,
sessment, prevention, and therapy of human cancers
tent: WO 0179556-A 959 25-OCT-2001;
llennium Predictive Medicine, Inc. (US)
Location/Qualifiers
1. .2041
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

39.0%; Score 30; DB 6; Length 2041;
imilarity 67.9%; Pred. No. 1.9;
; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 966

AGNACAACAGCCANCAGT 77
AGTACAACAGCCATCAGT 984

167367 2384 bp DNA linear PAT 17-DEC-2001
quence 3 from patent US 6287569.
167367
167367.1 GI:17903142

known.
known.
classified.
(bases 1 to 2384)
ps,T.J. and Wu,Y.
cines with enhanced intracellular processing
ent: US 6287569-A 3 11-SEP-2001;
Location/Qualifiers
1. .2384
/organism="unknown"
/mol_type="unassigned DNA"

39.0%; Score 30; DB 6; Length 2384;
milarity 67.9%; Pred. No. 1.9;
Conservative 0; Mismatches 24; Indels 1; Gaps 1;

CTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCATTTGGAGG 59
CTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCGATTGGAGG 1343

GNACAACAGCCANCAGT 77
GTACAACAGCCATCAGT 1361

466 2384 bp RNA linear PAT 29-SEP-1997
NA encoding tyrosinase.
466
466.1 GI:22025102
1996140699-A/1.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 2384)
bata,T., Suzuki,S., Takimoto,H. and Masui,S.
SUREMENT OF TYROSINASE MESSENGER RNA AMOUNT

JOURNAL Patent: JP 1996140699-A 1 04-JUN-1996;
POLA CHEM IND INC
COMMENT OS Homo sapiens (human)
PN JP 1996140699-A/1
PD 04-JUN-1996
PF 22-NOV-1994 JP 1994288041
PI SHIBATA TAKASHI, SUZUKI SATOSHI, TAKIMOTO HIROYUK
SHIGEKI
PC C12Q1/68,C12N15/09;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH
FT source 1. .2384
FT /organism='Homo sapiens'
FT /tissue_type='Skin'
FT /cell_type='melanocyte'
FT 5'UTR 1. .502
FT CDS 503. .2092
FT /product='tyrosinase'
FT 3'UTR 2093. .2384.
FEATURES Location/Qualifiers
source 1. .2384
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 2384;
Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
Db 1284 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77
Db 1344 AGTACAACAGCCATCAGT 1361

RESULT 31
I49609
LOCUS I49609 2384 bp DNA linear PA
DEFINITION Sequence 1 from patent US 5641508.
ACCESSION I49609
VERSION I49609.1 GI:2471829
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2384)
AUTHORS Li,L. and Lishko,V.K.
TITLE Method for delivering melanin to hair follicles
JOURNAL Patent: US 5641508-A 1 24-JUN-1997;
FEATURES Location/Qualifiers
source 1. .2384
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 39.0%; Score 30; DB 6; Length 2384;
Best Local Similarity 67.9%; Pred. No. 1.9;
Matches 53; Conservative 0; Mismatches 24; Indels 1;

Qy 1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
Db 1284 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy 60 AGNACAACAGCCANCAGT 77

```
GTACAACAGCCATCAGT 1361



ITYRA                  2384 bp    mRNA    linear   PRI 02-DEC-1996
ian tyrosinase (TYR) mRNA, complete cds.
'160
'160.1  GI:1698397
:osinase.
io sapiens (human)
io sapiens
:aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 601 to 2384)
.bahara,S., Tomita,Y., Tagami,H., Muller,R.M. and Cohen,T.
.ecular basis for the heterogeneity of human tyrosinase
ioku J. Exp. Med. 156 (4), 403-414 (1988)
22868
·4305
 (bases 1 to 600)
·eda,A., Tomita,Y., Okinaga,S., Tagami,H. and Shibahara,S.
.ctional analysis of the cDNA encoding human tyrosinase precursor
·chem. Biophys. Res. Commun. 162 (3), 984-990 (1989)
51001
4160
Dec 2, 1996 this sequence version replaced gi:340036.
ft entry and computer-readable copy of sequence [1] kindly
·vided by S.Shibahara, 15-AUG-1989.
      Location/Qualifiers
      1. .2384
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /chromosome="11"
      /map="11q14-q21"
      /cell_line="MeWo melanoma DNA, and cDNA to mRNA"
      /tissue_type="placenta"
      1. .2384
      /gene="TYR"
1     392. .399
      /gene="TYR"
      421. .2384
      /gene="TYR"
      /product="tyrosinase mRNA"
      503. .2092
      /gene="TYR"
      /EC_number="1.14.18.1"
      /codon_start=1
      /product="tyrosinase precursor"
      /protein_id="AAB37227.1"
      /db_xref="GI:340037"
      /db_xref="GDB:G00-120-476"
      /translation="MLLAVLYCLLWSFQTSAGHFPRACVSSKNLMEKECCPPWSGDRS
      PCGQLSGRGSCQNILLSNAPLGPQFPFTGVDDRESWPSVFYNRTCQCSGNFMGFNCGN
      CKFGFWGPNCTERRLLVRRNIFDLSAPEKDKFFAYLTLAKHTISSDYVIPIGTYGQMK
      NGSTPMFNDINIYDLFVWMHYYVSMDALLGGSEIWRDIDFAHEAPAFLPWHRLFLLRW
      EQEIQKLTGDENFTIPYWDWRDAEKCDICTDEYMGGQHPTNPNLLSPASFFSSWQIVC
      SRLEEYNSHQSLCNGTPEGPLRRNPGNHDKSRTPRLPSSADVEFCLSLTQYESGSMDK
      AANFSFRNTLEGFASPLTGIADASQSSMHNALHIYMNGTMSQVQGSANDPIFLLHHAF
      VDSIFEQWLRRHRPLQEVYPEANAPIGHNRESYMVPFIPLYRNGDFFISSKDLGYDYS
      YLQDSDPDSFQDYIKSYLEQASRIWSWLLGAAMVGAVLTALLAGLVSLLCRHKRKQLP
      EEKQPLLMEKEDYHSLYQSHL"
e     503. .556
      /gene="TYR"
e     557. .2089
      /gene="TYR"
      /product="tyrosinase"


          39.0%;  Score 30;  DB 9;  Length 2384;
milarity   67.9%;  Pred. No. 1.9;
  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATG-GTCAGGTCA
              |||||| |||||||||| ||||||       |   | |  ||  |  | ||
Db       1284 ACTTACTCAGCCCAGCATCATTCTTCTCCTCTTGGCAGATTGTCTGTAGCCG

Qy         60 AGNACAACAGCCANCAGT 77
              || |||||||||| ||||
Db       1344 AGTACAACAGCCATCAGT 1361


RESULT 33
AC121885/c
LOCUS       AC121885               155690 bp    DNA     linear   ROI
DEFINITION  Mus musculus BAC clone RP24-138N20 from chromosome 3, c
            sequence.
ACCESSION   AC121885
VERSION     AC121885.3  GI:22857828
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
REFERENCE   1  (bases 1 to 155690)
  AUTHORS   Shah,N., Cotton,M., Doebber,A. and Schatzkamer,K.
  TITLE     The sequence of Mus musculus BAC clone RP24-138N20
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 155690)
  AUTHORS   Wilson,R.
  TITLE     Sequencing of Mus musculus
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 155690)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAY-2002) Genome Sequencing Center, 4444
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 155690)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-2002) Genome Sequencing Center, 4444
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 155690)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-2002) Genome Sequencing Center, 4444
            Parkway, St. Louis, MO 63108, USA
REFERENCE   6  (bases 1 to 155690)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2003) Department of Genetics, Washinc
            University, 4444 Forest Park Avenue, St. Louis, Missour
COMMENT     On Sep 14, 2002 this sequence version replaced gi:22475
            -------------- Genome Center
                Center: Washington University Genome Sequencing Cen
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BB0138N20
            --------------.

            NOTICE:  This sequence may not represent the entire ins
            clone.  It may be shorter because we only sequence over
            clone sections once, or longer because we provide a sma
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise
            all regions were double stranded, sequenced with an alt
            chemistry, or covered by high quality data (i.e., phred
            30); an attempt was made to resolve all sequencing prob
            as compressions and repeats; all regions were covered b
            from more than one subclone; and the assembly was confi
            restriction digest.
```

PPING INFORMATION:
pping information for this clone was provided by Dr. Wes Warren,
partment of Genetics, Washington University, St. Louis MO. For
ditional information about the map position of this sequence, see
tp://genome.wustl.edu

JRCE INFORMATION:
e RPCI-24 BAC Library has been constructed by Pieter de Jong and
workers (http://www.chori.org) from male C57BL/6J mouse spleen
d/or brain genomic DNA. The clone and detailed information can be
tained from Pieter de Jong and coworkers at http://www.chori.org

IGHBORING SEQUENCE INFORMATION:
is sequence is the entire insert of the clone.

	Location/Qualifiers
	1. .155690
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="3"
	/map="3"
	/clone="RP24-138N20"
	/clone_lib="RPCI-24"
gion	221. .366
	/rpt_family="MaLR"
gion	389. .597
	/rpt_family="B2"
gion	612. .791
	/rpt_family="L1"
gion	992. .1278
	/rpt_family="MaLR"
gion	2544. .2646
	/rpt_family="Alu"
gion	3099. .3989
	/rpt_family="RMER13B"
gion	4291. .4371
	/rpt_family="B2"
gion	4380. .4589
	/rpt_family="B2"
gion	4628. .4967
	/rpt_family="ERVL"
gion	6288. .6485
	/rpt_family="B2"
gion	6574. .6885
	/rpt_family="MaLR"
gion	6886. .7362
	/rpt_family="ERVL"
gion	7848. .7942
	/rpt_family="B4"
gion	8007. .8079
	/rpt_family="B4"
gion	9216. .9370
	/rpt_family="B4"
gion	10005. .10075
	/rpt_family="ERVK"
gion	11588. .11859
	/rpt_family="B4"
gion	12238. .12631
	/rpt_family="MaLR"
gion	13019. .13179
	/rpt_family="B4"
gion	13528. .13800
	/rpt_family="B4"
gion	13927. .14064
	/rpt_family="Alu"
gion	16826. .18628
	/rpt_family="L1"
gion	20265. .20685
	/rpt_family="MaLR"
gion	20709. .20884
	/rpt_family="B2"
gion	20916. .21017
	/rpt_family="MaLR"
repeat_region	22202. .22485
	/rpt_family="MaLR"
repeat_region	23005. .23191
	/rpt_family="B2"
repeat_region	23270. .23333
	/rpt_family="ERV1"
repeat_region	23900. .24102
	/rpt_family="B2"
repeat_region	24278. .24333
	/rpt_family="B4"
repeat_region	24402. .24548
	/rpt_family="Alu"
repeat_region	24588. .24746
	/rpt_family="B4"
repeat_region	28996. .29273
	/rpt_family="MER2_type"
repeat_region	29274. .29704
	/rpt_family="RMER17C"
repeat_region	29705. .29783
	/rpt_family="MER2_type"
repeat_region	29903. .29980
	/rpt_family="MER2_type"
repeat_region	30871. .31011
	/rpt_family="Alu"
repeat_region	32984. .33155
	/rpt_family="B2"
repeat_region	33641. .33787
	/rpt_family="L1"
repeat_region	34591. .34640
	/rpt_family="B2"
repeat_region	34726. .34810
	/rpt_family="Alu"
repeat_region	37550. .37651
	/rpt_family="Alu"
repeat_region	39105. .39246
	/rpt_family="L1"
repeat_region	43851. .43916
	/rpt_family="L2"
repeat_region	44106. .44246
	/rpt_family="B2"
repeat_region	44793. .44986
	/rpt_family="B2"
repeat_region	46081. .46304
	/rpt_family="MER1_type"
repeat_region	46490. .46712
	/rpt_family="MIR"
repeat_region	47296. .47449
	/rpt_family="MaLR"
repeat_region	47605. .47719
	/rpt_family="MaLR"
repeat_region	48176. .48247
	/rpt_family="L2"
repeat_region	48830. .49151
	/rpt_family="ERV1"
repeat_region	49193. .49579
	/rpt_family="MaLR"
repeat_region	50485. .50687
	/rpt_family="B2"
repeat_region	53587. .53710
	/rpt_family="Alu"
repeat_region	55735. .56206
	/rpt_family="ERVL"
repeat_region	58819. .59027
	/rpt_family="L1"
repeat_region	59050. .59190
	/rpt_family="L1"

Query Match 37.4%; Score 28.8; DB 10; Length 1556
Best Local Similarity 62.7%; Pred. No. 6;
Matches 42; Conservative 0; Mismatches 25; Indels 0;

Qy 11 CCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATTTGGAGGAGN

```
CCCACTAACACTACAGCATGACCCTCTGCTCGCTCGGCTCTTCTGCACGAGAACACTCCC 88239

CANCAGT 77
   ||||
TGCCAGT 88232



118220              162383 bp    DNA     linear   HTG 28-AUG-2002
s musculus clone RP23-272M11, WORKING DRAFT SEQUENCE, 2 unordered
eces.
118220
118220.3  GI:22532661
3; HTGS_PHASE1; HTGS_DRAFT.
s musculus (house mouse)
s musculus
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 (bases 1 to 162383)
rren,B., Nusbaum,C. and Lander,E.
s musculus, clone RP23-272M11
ublished
 (bases 1 to 162383)
rren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
derson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
ukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
azaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
ok,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
co,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
ude,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
gos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
mat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
nders,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
:Lean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
arthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
ova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
bu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
ver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
mond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
an,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
man,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
auss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
ham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
l,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
noun,J., Zembek,L., Zimmer,A. and Zody,M.
ect Submission
mitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
earch, 320 Charles Street, Cambridge, MA 02141, USA
(bases 1 to 162383)
ren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
na,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
arata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
k,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
o,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
dyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
ton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
atas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
arthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
phy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
onnor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
nkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
an,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
th,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
faye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
l,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
bek,L., Zimmer,A. and Zody,M.
ect Submission
mitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome
earch, 320 Charles Street, Cambridge, MA 02141, USA
Aug 28, 2002 this sequence version replaced gi:22267708.
 repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            ------------- Project Information
                Center project name: L23645
                Center clone name: 272_M_11
            ------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 162137 bases at least Q40
                Consensus quality: 162160 bases at least Q30
                Consensus quality: 162212 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 162283; sum-of-contigs
                Quality coverage: 12.2 in Q20 bases; agarose-fp
                Quality coverage: 12.3 in Q20 bases; sum-of-contig
            -------------.
            * NOTE: This is a 'working draft' sequence. It current
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record
            * arbitrary. Gaps between the contigs are represented
            * runs of N, but the exact sizes of the gaps are unkno
            * This record will be updated with the finished sequen
            * as soon as it is available and the accession number
            * be preserved.
            *        1   160172: contig of 160172 bp in length
            *   160173   160272: gap of 100 bp
            *   160273   162383: contig of 2111 bp in length.
FEATURES             Location/Qualifiers
     source          1. .162383
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-272M11"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .160172
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    160273. .162383
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             37.1%;  Score 28.6;  DB 2;  Length 162383
  Best Local Similarity   65.6%;  Pred. No. 7.2;
  Matches   40;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 ACTTACNCAGCCCAGCANCATTCTAGCATCATCCTCTGCATGGTCAGGTCATT
              ||| |  || |  || |    ||||||| ||| | |||||   | ||  ||||
Db      27361 ACTAAGCCAACTGAGTAAATATCTAGCAGCATTCACTGCAGCTTGAGTCCATT

Qy         61 G 61
              |
Db      27301 G 27301


RESULT 35
BD131906
LOCUS       BD131906                  67 bp    DNA     linear   PAT
DEFINITION  Novel DNA cloning method.
ACCESSION   BD131906
VERSION     BD131906.1  GI:23226851
KEYWORDS    JP 2002503448-A/24.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
```

(bases 1 to 67)
ewart,F., Zhang,Y. and Buchholz,F.
vel DNA cloning method
tent: JP 2002503448-A 24 05-FEB-2002;
ROPAEISCHES LABORATORIUM FUER MOLEKULARBIOLOGIE
Artificial Sequence
JP 2002503448-A/24
05-FEB-2002
07-DEC-1998 JP 2000524410
05-DEC-1997 EP 97121462.2,05-OCT-1998 EP 98118756.0 PI
ANCIS STEWART,YOUMING ZHANG,FRANK BUCHHOLZ
C12N15/09//(C12N15/09,C12R1:19),C12N15/00,(C12N15/00,C12R1:19)
Description of Artificial Sequence: Artificial Sequence FH
y Location/Qualifiers
source 1. .67
/organism='Artificial Sequence'.
Location/Qualifiers
1. .67
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

36.9%; Score 28.4; DB 6; Length 67;
imilarity 96.7%; Pred. No. 7.4;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
CGAGCATCATCCTCTGCATGGTCAGGTCAT 32

131908 67 bp DNA linear PAT 18-SEP-2002
vel DNA cloning method.
131908
131908.1 GI:23226853
2002503448-A/26.
athetic construct
athetic construct
tificial sequences.
(bases 1 to 67)
ewart,F., Zhang,Y. and Buchholz,F.
vel DNA cloning method
tent: JP 2002503448-A 26 05-FEB-2002;
ROPAEISCHES LABORATORIUM FUER MOLEKULARBIOLOGIE
Artificial Sequence
JP 2002503448-A/26
05-FEB-2002
07-DEC-1998 JP 2000524410
05-DEC-1997 EP 97121462.2,05-OCT-1998 EP 98118756.0 PI
ANCIS STEWART,YOUMING ZHANG,FRANK BUCHHOLZ
C12N15/09//(C12N15/09,C12R1:19),C12N15/00,(C12N15/00,C12R1:19)
Description of Artificial Sequence: Artificial Sequence FH
Location/Qualifiers
source 1. .67
/organism='Artificial Sequence'.
Location/Qualifiers
1. .67
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

36.9%; Score 28.4; DB 6; Length 67;
imilarity 96.7%; Pred. No. 7.4;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
CGAGCATCATCCTCTGCATGGTCAGGTCAT 40

RESULT 37
AB128155
LOCUS AB128155 299 bp mRNA linear MA
DEFINITION Oryctolagus cuniculus Lac Z mRNA for beta-D-galactosid
cds.
ACCESSION AB128155
VERSION AB128155.1 GI:39725589
KEYWORDS .
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus
REFERENCE 1
AUTHORS Murata,T.
TITLE Vascular endothelium has a local anti-adenovirus vecto
glucocorticoid optimizes its gene transduction
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 299)
AUTHORS Murata,T.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2003) Takahisa Murata, The Universit
Department of Veterinary Pharmacology, Graduate School
Agriculture and Life Sciences; 1-1-1 Yayoi, Bunkyo-ku,
113-8657, Japan (E-mail:murata@mail.vm.a.u-tokyo.ac.jp
Tel:81-3-5841-5394, Fax:81-3-5841-8183)
FEATURES Location/Qualifiers
source 1. .299
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/db_xref="taxon:9986"
gene 1. .299
/gene="Lac Z"
CDS <1. .>299
/gene="Lac Z"
/codon_start=1
/product="beta-D-galactosidase"
/protein_id="BAD04918.1"
/db_xref="GI:39725590"
/translation="VNRHEHHPLHGQVMDEQTMVQDILLMKQNNF
WYTLCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTR
ORIGIN

Query Match 36.9%; Score 28.4; DB 4; Length 299;
Best Local Similarity 96.7%; Pred. No. 7.6;
Matches 29; Conservative 0; Mismatches 1; Indels 0

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
Db 12 CGAGCATCATCCTCTGCATGGTCAGGTCAT 41

RESULT 38
ECO487615
LOCUS ECO487615 1063 bp DNA linear BC
DEFINITION Escherichia coli partial lacZ gene for beta-D-galactos
mutant 645.
ACCESSION AJ487615
VERSION AJ487615.1 GI:21104348
KEYWORDS beta-D-galactosidase; lacZ gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterob
Enterobacteriaceae; Escherichia.
REFERENCE 1
AUTHORS Laden,J.C., Philibert,P., Torreilles,F., Pugniere,M. ar
Martineau,P.
TITLE Expression and folding of an antibody fragment selected
high expression levels in Escherichia coli cytoplasm
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1063)
AUTHORS Martineau,P.
TITLE Direct Submission

omitted (06-MAY-2002) Martineau P., UMR5094 Faculte de Pharmacie,
rs, 15 av. Charles Flahault BP14491, 34093 Montpellier Cedex 5,
ANCE
Location/Qualifiers
1. .1063
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="AMEF mutant 645"
1. .1063
/gene="lacZ"
<1. .>1063
/gene="lacZ"
/codon_start=2
/evidence=experimental
/transl_table=11
/product="beta-D-galactosidase"
/protein_id="CAD32242.1"
/db_xref="GI:21104349"
/db_xref="GOA:Q8L1L8"
/db_xref="SPTREMBL:Q8L1L8"
/translation="GSYLEDQDMWRMSDIFRDVSLLHKPTTQISDFHVATRFNDDFSR
AVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGYADRVTLRLNV
ENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLIRG
VNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYTLCDRYGLYVVD
EANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHD
ALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVPKWSIKKWLSLP
GETRPLILCEYAHAMGNSLG"
42
/gene="lacZ"
/note="wild-type"
/replace="g"

36.9%; Score 28.4; DB 1; Length 1063;
milarity 96.7%; Pred. No. 7.8;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

TAGCATCATCCTCTGCATGGTCAGGTCAT 52
||||||||||||||||||||||||||||
GAGCATCATCCTCTGCATGGTCAGGTCAT 522

)487616 1063 bp DNA linear BCT 21-MAY-2002
herichia coli partial lacZ gene for beta-D-galactosidase, AMEF
ant 959.
87616
87616.1 GI:21104350
a-D-galactosidase; lacZ gene.
herichia coli
herichia coli
teria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
erobacteriaceae; Escherichia.

en,J.C., Philibert,P., Torreilles,F., Pugniere,M. and
tineau,P.
ression and folding of an antibody fragment selected in vivo for
h expression levels in Escherichia coli cytoplasm
ublished
(bases 1 to 1063)
tineau,P.
ect Submission
mitted (06-MAY-2002) Martineau P., UMR5094 Faculte de Pharmacie,
s, 15 av. Charles Flahault BP14491, 34093 Montpellier Cedex 5,
NCE
Location/Qualifiers
1. .1063
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="AMEF mutant 959"
gene 1. .1063
/gene="lacZ"
CDS <1. .>1063
/gene="lacZ"
/codon_start=2
/evidence=experimental
/transl_table=11
/product="beta-D-galactosidase"
/protein_id="CAD32243.1"
/db_xref="GI:21104351"
/db_xref="GOA:Q8KHF9"
/db_xref="SPTREMBL:Q8KHF9"
/translation="GSYLEDQDMWRMSGIFRDVSLLHKPTTQISD
AVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDER
ENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENG
VNRHKHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLW
EANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIW
ALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPA
GETRPLILCEYAHAMGNSLG"
variation 494
/gene="lacZ"
/note="wild-type"
/replace="g"
ORIGIN

Query Match 36.9%; Score 28.4; DB 1; Length 1063;
Best Local Similarity 96.7%; Pred. No. 7.8;
Matches 29; Conservative 0; Mismatches 1; Indels 0

Qy 23 CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||
Db 493 CAAGCATCATCCTCTGCATGGTCAGGTCAT 522

RESULT 40
ECO487617
LOCUS ECO487617 1063 bp DNA linear BC
DEFINITION Escherichia coli partial lacZ201 gene for beta-D-galact
AMEF mutant W6106.
ACCESSION AJ487617
VERSION AJ487617.1 GI:21104352
KEYWORDS beta-D-galactosidase; lacZ201 gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterob
Enterobacteriaceae; Escherichia.
REFERENCE 1
AUTHORS Laden,J.C., Philibert,P., Torreilles,F., Pugniere,M. a
Martineau,P.
TITLE Expression and folding of an antibody fragment selected
high expression levels in Escherichia coli cytoplasm
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1063)
AUTHORS Martineau,P.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Martineau P., UMR5094 Faculte d
Cnrs, 15 av. Charles Flahault BP14491, 34093 Montpellie
FRANCE
FEATURES Location/Qualifiers
source 1. .1063
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="AMEF mutant W6106"
gene 1. .1063
/gene="lacZ201"
CDS <1. .>1063
/gene="lacZ201"
/codon_start=2
/evidence=experimental
/transl_table=11
/product="beta-D-galactosidase"
/protein_id="CAD32244.1"

/db_xref="GI:21104353"
/db_xref="GOA:Q8KHF9"
/db_xref="SPTREMBL:Q8KHF9"
/translation="GSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFHVATRFNDDFSR
AVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGYADRVTLRLNV
ENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLIRG
VNRHKHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYTLCDRYGLYVVD
EANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHD
ALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVPKWSIKKWLSLP
GETRPLILCEYAHAMGNSLG"
494
/gene="lacZ201"
/note="wild-type"
/replace="g"

36.9%; Score 28.4; DB 1; Length 1063;
imilarity 96.7%; Pred. No. 7.8;
; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CTAGCATCATCCTCTGCATGGTCAGGTCAT 52
| ||||||||||||||||||||||||||||||
CAAGCATCATCCTCTGCATGGTCAGGTCAT 522

ed: April 13, 2004, 19:12:11
9.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

ıcleic search, using sw model

April 13, 2004, 18:17:02 ; Search time 1853.5 Seconds
(without alignments)
1240.565 Million cell updates/sec

US-10-090-326-24
77
1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

27513289 seqs, 14931090276 residues

hits satisfying chosen parameters: 55026578

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

is the number of results predicted by chance to have a
ater than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

	% Query Match	Length	DB	ID	Description
	58.4	383	12	BM752895	BM752895 K-EST0029
	58.4	412	10	BE696948	BE696948 RC1-CT040
	56.4	196	10	BF755519	BF755519 MR2-CT059
	56.4	237	12	BG950112	BG950112 CM2-CT063
	43.4	284	10	BF818092	BF8180
	43.4	295	10	BE697671	BE6976
c	43.4	302	12	BG957601	BG9576
	43.4	401	12	BG981406	BG9814
c	43.4	417	12	BG955581	BG9555
c	43.4	524	10	AW362544	AW3625
	43.4	540	10	AW937852	AW9378
	43.4	556	10	BF759065	BF7590
	43.4	556	12	BG956260	BG9562
c	43.4	563	10	BF805557	BF8055
	43.4	579	10	AW604315	AW6043
	43.4	593	12	BM742895	BM7428
	43.4	596	10	AW858762	AW8587
	43.4	601	12	BM786353	BM7863
c	43.4	612	10	AW360942	AW3609
c	43.4	644	10	AW853522	AW8535
	43.4	904	13	BQ685640	BQ6856
	43.4	907	13	BQ689619	BQ6896
c	41.8	372	10	AW360989	AW3609
	41.8	402	10	BF756277	BF7562
	41.8	444	12	BM836580	BM8365
	41.6	423	9	AA055605	AA05560
	41	312	9	AA366576	AA36657
	41	389	9	AA132613	AA13261
c	40.2	287	10	BF817244	BF8172
	40.2	327	9	AA134992	AA13499
	40.2	630	13	BX480185	BX4801
	38.4	494	10	BF084626	BF0846
	38.4	623	10	AW862017	AW8620
	37.6	440	9	AA127122	AA12712
	37	541	10	AW937770	AW9377
	35.2	590	10	AW361304	AW3613
c	35	296	12	BG950148	BG9501
	34.6	914	10	BE735665	BE7356
	34.4	237	9	AA327327	AA32732
	34.4	423	10	AW849404	AW8494
	33.6	338	10	AW383654	AW3836
	33.2	577	13	BX644443	BX6444
c	32.8	233	10	AW383495	AW3834
	32.8	301	12	BG950114	BG9501
	32.8	321	10	BF336900	BF3369

ALIGNMENTS

RESULT 1
BM752895
LOCUS BM752895 383 bp mRNA linear EST
DEFINITION K-EST0029613 S9SNU601 Homo sapiens cDNA clone S9SNU601-
mRNA sequence.
ACCESSION BM752895
VERSION BM752895.1 GI:19082513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 383)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., K
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 8 row: G column: 11
High quality sequence stop: 383.

Location/Qualifiers
1. .383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-8-G11"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F'"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

58.4%; Score 45; DB 12; Length 383;
imilarity 74.0%; Pred. No. 0.00093;
; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTAGCCAAGCCCTCCATCTCCAGCAAAAACTCCAAAC 337

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 354

696948 412 bp mRNA linear EST 11-SEP-2000
1-CT0408-220700-011-f06 CT0408 Homo sapiens cDNA, mRNA sequence.
696948
696948.1 GI:10084108
P.
no sapiens (human)
no sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 412)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
npson,A.J.
otgun sequencing of the human transcriptome with ORF expressed
quence tags
oc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
202663
737800
ntact: Simpson A.J.G.
oratory of Cancer Genetics
dwig Institute for Cancer Research
a Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
azil
l: +55-11-2704922
x: +55-11-2707001
ail: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human C
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=R
700-011-f06&t3=2000-07-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 412.
FEATURES Location/Qualifiers
source 1. .412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0408"
/note="Organ: colon; Vector: puc18; Site_1: S
SmaI; A mini-library was made by cloning prod
from ORESTES PCR (U.S. Letters Patent applic
196,716 - Ludwig Institute for Cancer Resear
into the pUC 18 vector. Reverse transcriptio
mRNA and cDNA amplification were performed un
stringency conditions."
ORIGIN

Query Match 58.4%; Score 45; DB 10; Length 412;
Best Local Similarity 74.0%; Pred. No. 0.00096;
Matches 57; Conservative 0; Mismatches 20; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||| | | | | || | | ||
Db 331 AGACAATCACAGTCTCTGCGGAGCTGCCCTAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 271 CCGTGGAGGACAAGGAT 255

RESULT 3
BF755519/c
LOCUS BF755519 196 bp mRNA linear ES
DEFINITION MR2-CT0595-311000-002-f10 CT0595 Homo sapiens cDNA, mR
ACCESSION BF755519
VERSION BF755519.1 GI:12103419
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE 1 (bases 1 to 196)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cos
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de S
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sa
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human C
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR2&t
311000-002-f10&t3=2000-10-31&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 196.

Location/Qualifiers
1. .196
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0595"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

56.4%; Score 43.4; DB 10; Length 196;
imilarity 72.7%; Pred. No. 0.0022;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 116

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 99

)50112 237 bp mRNA linear EST 12-JUN-2001
?-CT0639-030101-681-a04 CT0639 Homo sapiens cDNA, mRNA sequence.
'50112
'50112.1 GI:14368283
'.
io sapiens (human)
io sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 237)
s Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
ai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
dman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
are,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
pson,A.J.
tgun sequencing of the human transcriptome with ORF expressed
uence tags
c. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-CT0639-
101-681-a04&t3=2001-01-03&t4=1)
primer: puc 18 forward
h quality sequence start: 12
h quality sequence stop: 237.
Location/Qualifiers
1. .237
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0639"
/note="Organ: colon; Vector: puc18; Site_1: S
SmaI; A mini-library was made by cloning prod
from ORESTES PCR (U.S. Letters Patent applic
196,716 - Ludwig Institute for Cancer Resear
into the pUC 18 vector. Reverse transcriptic
mRNA and cDNA amplification were performed un
stringency conditions."
ORIGIN

Query Match 56.4%; Score 43.4; DB 12; Length 237;
Best Local Similarity 72.7%; Pred. No. 0.0024;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| | | | || | | ||
Db 93 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 153 CCGTGGAGGACAAGGAT 169

RESULT 5
BF818092
LOCUS BF818092 284 bp mRNA linear ES
DEFINITION MR3-CI0184-151200-007-c04 CI0184 Homo sapiens cDNA, mRI
ACCESSION BF818092
VERSION BF818092.1 GI:12155181
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE 1 (bases 1 to 284)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., I
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., S
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Sc
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sac
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2
151200-007-c04&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 284.
FEATURES Location/Qualifiers
source 1. .284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CI0184"
/note="Organ: colon_ins; Vector: puc18; Site_1
Site_2: SmaI; A mini-library was made by cloni
derived from ORESTES PCR (U.S. Letters Patent
No. 196,716 - Ludwig Institute for Cancer Res
profiles into the pUC 18 vector. Reverse tran
tissue mRNA and cDNA amplification were perfor

low stringency conditions."

56.4%; Score 43.4; DB 10; Length 284;
imilarity 72.7%; Pred. No. 0.0026;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||  |   |  ||  |  | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 143

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 160
```

697671 295 bp mRNA linear EST 11-SEP-2000
0-CT0471-080800-032-e09 CT0471 Homo sapiens cDNA, mRNA sequence.
697671
697671.1 GI:10084935
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 295)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
npson,A.J.
otgun sequencing of the human transcriptome with ORF expressed
quence tags
oc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
202663
737800
ntact: Simpson A.J.G.
boratory of Cancer Genetics
dwig Institute for Cancer Research
a Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
azil
l: +55-11-2704922
x: +55-11-2707001
ail: asimpson@ludwig.org.br
is sequence was derived from the FAPESP/LICR Human Cancer Genome
oject. This entry can be seen in the following URL
ttp://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC0-CT0471-080
0-032-e09&t3=2000-08-08&t4=1)
g primer: puc 18 forward
gh quality sequence start: 21
gh quality sequence stop: 295.
Location/Qualifiers
1. .295
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0471"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

56.4%; Score 43.4; DB 10; Length 295;
imilarity 72.7%; Pred. No. 0.0027;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
              ||||||||||||||||||||||  |   |  ||  |  | ||
Db         75 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        135 CCGTGGAGGACAAGGAT 151
```

RESULT 7
BG957601/c
LOCUS BG957601 302 bp mRNA linear ES
DEFINITION PM0-CT0802-020301-001-e06 CT0802 Homo sapiens cDNA, mR
ACCESSION BG957601
VERSION BG957601.1 GI:14375772
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE 1 (bases 1 to 302)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cos
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de S
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sa
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human C
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t
020301-001-e06&t3=2001-03-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 247.
FEATURES Location/Qualifiers
source 1. .302
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0802"
/note="Organ: colon; Vector: puc18; Site_1: S
SmaI; A mini-library was made by cloning prod
from ORESTES PCR (U.S. Letters Patent applic
196,716 - Ludwig Institute for Cancer Resear
into the pUC 18 vector. Reverse transcriptio
mRNA and cDNA amplification were performed un
stringency conditions."
ORIGIN

Query Match 56.4%; Score 43.4; DB 12; Length 302;
Best Local Similarity 72.7%; Pred. No. 0.0027;
Matches 56; Conservative 0; Mismatches 21; Indels 0,

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
              ||||||||||||||||||||||  |   |  ||  |  | ||
Db        227 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy         61 CCGTGGAGGACAAGGAT 77
```

```
||||||||||||||||
'GTGGAGGACAAGGAT 151



:1406                    401 bp    mRNA    linear   EST 12-JUN-2001
'CN0101-070201-470-G12 CN0101 Homo sapiens cDNA, mRNA sequence.
)1406
)1406.1  GI:14384141

) sapiens (human)
) sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
nalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 401)
s Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
ai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
iman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
are,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
)son,A.J.
:gun sequencing of the human transcriptome with ORF expressed
uence tags
:. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
)2663
)7800
:act: Simpson A.J.G.
)ratory of Cancer Genetics
vig Institute for Cancer Research
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-CN0101-
201-470-G12&t3=2001-02-07&t4=1)
 primer: puc 18 forward
h quality sequence stop: 401.
      Location/Qualifiers
      1. .401
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="CN0101"
      /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
      Site_2: SmaI; A mini-library was made by cloning products
      derived from  ORESTES PCR (U.S. Letters Patent application
      No. 196,716 -  Ludwig Institute for Cancer Research)
      profiles into  the pUC 18 vector. Reverse transcription of
      tissue mRNA and cDNA amplification were performed under
      low stringency conditions."


          56.4%;  Score 43.4;  DB 12;  Length 401;
milarity   72.7%;  Pred. No. 0.003;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||      |     |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 191

'CGTGGAGGACAAGGAT 77
||||||||||||||||
'CGTGGAGGACAAGGAT 208



)55581                   417 bp    mRNA    linear   EST 12-JUN-2001
```

```
DEFINITION  CM4-CT0656-130201-852-c02 CT0656 Homo sapiens cDNA, mRN
ACCESSION   BG955581
VERSION     BG955581.1  GI:14373752
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1  (bases 1 to 417)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., B
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., S
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de So
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Ca
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM4&t2
            130201-852-c02&t3=2001-02-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 34
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0656"
                     /note="Organ: colon; Vector: puc18; Site_1: Sm
                     SmaI; A mini-library was made by cloning produ
                     from  ORESTES PCR (U.S. Letters Patent applica
                     196,716 -  Ludwig Institute for Cancer Researc
                     into  the pUC 18 vector. Reverse transcription
                     mRNA and cDNA amplification were performed und
                     stringency conditions."
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 12;  Length 417;
  Best Local Similarity   72.7%;  Pred. No. 0.0031;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||      |     |  |    ||  |  | |||
Db        367 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              ||||||||||||||||
Db        307 CCGTGGAGGACAAGGAT 291


RESULT 10
AW362544/c
LOCUS       AW362544                 524 bp    mRNA    linear   EST
DEFINITION  RC3-CT0283-201199-011-g06 CT0283 Homo sapiens cDNA, mRN
ACCESSION   AW362544
VERSION     AW362544.1  GI:6867194
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 524)
P http://www.ludwig.org.br/ORESTES.
FAPESP/LICR Human Cancer Genome Project
ublished (1999)
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC3&t2=RC3-CT0283-
199-011-g06&t3=1999-11-20&t4=1)
primer: puc 18 forward
h quality sequence start: 12
h quality sequence stop: 524.
Location/Qualifiers
1. .524
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0283"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

56.4%; Score 43.4; DB 10; Length 524;
milarity 72.7%; Pred. No. 0.0034;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

.GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||| | | || | | ||||||||||
.GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 202

:CGTGGAGGACAAGGAT 77
|||||||||||||||
:CGTGGAGGACAAGGAT 185

)37852 540 bp mRNA linear EST 30-MAY-2000
)-DT0045-010300-108-f02 DT0045 Homo sapiens cDNA, mRNA sequence.
)37852
)37852.1 GI:8113281
F.
no sapiens (human)
no sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 540)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
instein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
mpson,A.J.
otgun sequencing of the human transcriptome with ORF expressed
quence tags
oc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
202663
737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV
300-1C8-f02&t3=2000-03-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 9
High quality sequence stop: 540.
FEATURES Location/Qualifiers
source 1. .540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="DT0045"
/note="Organ: denis_drash; Vector: puc18; Site
Site_2: SmaI; A mini-library was made by cloni
derived from ORESTES PCR (U.S. Letters Patent
No. 196,716 - Ludwig Institute for Cancer Res
profiles into the pUC 18 vector. Reverse tran
tissue mRNA and cDNA amplification were perfor
low stringency conditions."
ORIGIN

Query Match 56.4%; Score 43.4; DB 10; Length 540;
Best Local Similarity 72.7%; Pred. No. 0.0034;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||| | | || | | |||
Db 378 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 438 CCGTGGAGGACAAGGAT 454

RESULT 12
BF759065
LOCUS BF759065 556 bp mRNA linear EST
DEFINITION QV4-CT0069-281100-604-h04 CT0069 Homo sapiens cDNA, mRN
ACCESSION BF759065
VERSION BF759065.1 GI:12106965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 556)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., E
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., S
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de So
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922

: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV4&t2=QV4-CT0069-
100-604-h04&t3=2000-11-28&t4=1)
primer: puc 18 forward
h quality sequence start: 223
h quality sequence stop: 555.
Location/Qualifiers
1. .556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0069"
/note="Organ: colon; Vector: pUC18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

56.4%; Score 43.4; DB 10; Length 556;
milarity 72.7%; Pred. No. 0.0035;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |     | |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 422

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 439
```

56260 556 bp mRNA linear EST 12-JUN-2001
-CT0616-190201-553-b02 CT0616 Homo sapiens cDNA, mRNA sequence.
56260
56260.1 GI:14374431
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 556)
s Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
ai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
dman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
are,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
pson,A.J.
tgun sequencing of the human transcriptome with ORF expressed
uence tags
c. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-CT0616-
201-553-b02&t3=2001-02-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 549.
FEATURES Location/Qualifiers
source 1. .556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0616"
/note="Organ: colon; Vector: puc18; Site_1: Sm
SmaI; A mini-library was made by cloning produ
from ORESTES PCR (U.S. Letters Patent applica
196,716 - Ludwig Institute for Cancer Researc
into the pUC 18 vector. Reverse transcription
mRNA and cDNA amplification were performed und
stringency conditions."
ORIGIN

Query Match 56.4%; Score 43.4; DB 12; Length 556;
Best Local Similarity 72.7%; Pred. No. 0.0035;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||     |     | |    ||  |  | |||
Db        284 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        344 CCGTGGAGGACAAGGAT 360
```

RESULT 14
BF805557/c
LOCUS BF805557 563 bp mRNA linear EST
DEFINITION QV1-CI0173-061100-462-g02 CI0173 Homo sapiens cDNA, mRN
ACCESSION BF805557
VERSION BF805557.1 GI:12134546
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 563)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., B
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., S
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de So
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV1&t2
061100-462-g02&t3=2000-11-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 563.
FEATURES Location/Qualifiers
source 1. .563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/dev_stage="Adult"
/clone_lib="CI0173"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

56.4%; Score 43.4; DB 10; Length 563;
milarity 72.7%; Pred. No. 0.0035;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 162

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 145

04315 579 bp mRNA linear EST 23-MAR-2000
-CT0219-280100-062-G01 CT0219 Homo sapiens cDNA, mRNA sequence.
04315
04315.1 GI:7309056
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 579)
P http://www.ludwig.org.br/ORESTES.
: FAPESP/LICR Human Cancer Genome Project
ublished (1999)
itact: Simpson A.J.G.
oratory of Cancer Genetics
dwig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
.: +55-11-2704922
c: +55-11-2707001
iil: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
oject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-CT0219-
0100-062-G01&t3=2000-01-28&t4=1)
primer: puc 18 forward
gh quality sequence stop: 572.
Location/Qualifiers
1. .579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0219"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

56.4%; Score 43.4; DB 10; Length 579;
imilarity 72.7%; Pred. No. 0.0035;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 471 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
||||||||||||||||
Db 531 CCGTGGAGGACAAGGAT 547

RESULT 16
BM742895
LOCUS BM742895 593 bp mRNA linear EST
DEFINITION K-EST0015932 S6SNU620 Homo sapiens cDNA clone S6SNU620-mRNA sequence.
ACCESSION BM742895
VERSION BM742895.1 GI:19064224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 593)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., K
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 10 row: G column: 05
High quality sequence stop: 593.
FEATURES Location/Qualifiers
source 1. .593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S6SNU620-10-G05"
/sex="F"
/tissue_type="Ascites"
/cell_type="Scattering floating"
/cell_line="SNU-620"
/lab_host="Top10F'"
/clone_lib="S6SNU620"
/note="Organ: Stomach; Vector: pCNS; Site_1: I
Site_2: NotI; The poly (A)+ RNA was dephospho
bacterial alkaline phosphatase (BAP) and then
with tabacco acid pyrophosphatase (TAP). The
intact mRNA was ligated with DNA-RNA linker i
I site by treatment of T4 RNA ligase and the
cDNA was synthesized from oligo dT-selected m
priming with dT-tailed vector. The dT-tailed
adjusted to have about 60nt. The cDNA vector
circularized with E. coli DNA ligase after di
EcoRI which site is also included in vector.
converted to a DNA strand by Okayama-Berg met
obtained cDNA vectors were used for transform
competent cells E. coli Top10F' by electropor
The cDNA libraries constructed by this method
full-length enriched cDNA library."
ORIGIN

Query Match 56.4%; Score 43.4; DB 12; Length 593;
Best Local Similarity 72.7%; Pred. No. 0.0036;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| | | | || | | ||
Db 240 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

```
GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 316
```

```
8762                    596 bp    mRNA    linear   EST 19-MAY-2000
CT0347-160200-013-h10 CT0347 Homo sapiens cDNA, mRNA sequence.
8762
8762.1  GI:7954455

 sapiens (human)
 sapiens
ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
bases 1 to 596)
 Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
i,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
man,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
stein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
re,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
son,A.J.
gun sequencing of the human transcriptome with ORF expressed
ence tags
. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2663
7800
act: Simpson A.J.G.
ratory of Cancer Genetics
ig Institute for Cancer Research
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
il
 +55-11-2704922
 +55-11-2707001
1: asimpson@ludwig.org.br
 sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
p://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-CT0347-160
013-h10&t3=2000-02-16&t4=1)
 primer: puc 18 forward
 quality sequence stop: 593.
      Location/Qualifiers
      1. .596
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="CT0347"
      /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
      SmaI; A mini-library was made by cloning products derived
      from  ORESTES PCR (U.S. Letters Patent application No.
      196,716 -  Ludwig Institute for Cancer Research) profiles
      into  the pUC 18 vector. Reverse transcription of tissue
      mRNA and cDNA amplification were performed under low
      stringency conditions."


          56.4%;  Score 43.4;  DB 10;  Length 596;
nilarity   72.7%;  Pred. No. 0.0036;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||     |   |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 154

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 171
```

```
BM786353                                                        -
LOCUS       BM786353                 601 bp    mRNA    linear   EST
DEFINITION  K-EST0065059 S9SNU601 Homo sapiens cDNA clone S9SNU601-3
            mRNA sequence.
ACCESSION   BM786353
VERSION     BM786353.1  GI:19134585
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y.
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Ki
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: E  column: 10
            High quality sequence stop: 601.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-36-E10"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_
                     Site_2: XhoI; The poly (A)+ RNA was dephosphor
                     bacterial alkaline phosphatase (BAP) and then
                     with tabacco acid pyrophosphatase (TAP). The d
                     intact mRNA was ligated with DNA-RNA linker in
                     site by treatment of T4 RNA ligase and the fir
                     cDNA was synthesized with Superscript II using
                     oligo-dT primer. After first strand synthesis,
                     degraded by NaOH treatment and cDNA was amplif
                     reaction. The PCR products were digested with
                     cloned into DraIII- digested pME18S-FL3 vector
                     obtained cDNA vectors were used for transforma
                     competent cells E. coli Top10F' by electropora
                     The cDNA libraries constructed by this method
                     full-length enriched cDNA library."
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 12;  Length 601;
  Best Local Similarity   72.7%;  Pred. No. 0.0036;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||     |   |  |    ||  |  | |||
Db        500 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        560 CCGTGGAGGACAAGGAT 576


RESULT 19
AW360942/c
LOCUS       AW360942                 612 bp    mRNA    linear   EST
DEFINITION  PM1-CT0247-151299-005-e09 CT0247 Homo sapiens cDNA, mRN
ACCESSION   AW360942
```

```
60942.1  GI:6865592
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 612)
P http://www.ludwig.org.br/ORESTES.
 FAPESP/LICR Human Cancer Genome Project
ublished (1999)
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-CT0247-
299-005-e09&t3=1999-12-15&t4=1)
 primer: puc 18 forward
h quality sequence start: 24
h quality sequence stop: 598.
      Location/Qualifiers
      1. .612
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="CT0247"
      /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
      SmaI; A mini-library was made by cloning products derived
      from  ORESTES PCR (U.S. Letters Patent application No.
      196,716 -  Ludwig Institute for Cancer Research) profiles
      into  the pUC 18 vector. Reverse transcription of tissue
      mRNA and cDNA amplification were performed under low
      stringency conditions."


          56.4%;  Score 43.4;  DB 10;  Length 612;
milarity   72.7%;  Pred. No. 0.0036;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |     |   |   ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 171

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 154



)53522                  644 bp    mRNA    linear   EST 19-MAY-2000
.-CT0252-140300-026-g06 CT0252 Homo sapiens cDNA, mRNA sequence.
)53522
)53522.1  GI:7949267
.
no sapiens (human)
no sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 644)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
instein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
mpson,A.J.
otgun sequencing of the human transcriptome with ORF expressed
          sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sac
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Ca
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC
          300-026-g06&t3=2000-03-14&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 24
          High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0252"
                     /note="Organ: colon; Vector: puc18; Site_1: Sm
                     SmaI; A mini-library was made by cloning produ
                     from  ORESTES PCR (U.S. Letters Patent applica
                     196,716 -  Ludwig Institute for Cancer Researc
                     into  the pUC 18 vector. Reverse transcriptior
                     mRNA and cDNA amplification were performed unc
                     stringency conditions."
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 10;  Length 644;
  Best Local Similarity   72.7%;  Pred. No. 0.0037;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||| ||||||||||||||||||     |     |   |   ||  |  | |||
Db        360 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        300 CCGTGGAGGACAAGGAT 284


RESULT 21
BQ685640
LOCUS       BQ685640                904 bp    mRNA    linear   EST
DEFINITION  AGENCOURT_8345228 NIH_MGC_110 Homo sapiens cDNA clone I
            5', mRNA sequence.
ACCESSION   BQ685640
VERSION     BQ685640.1  GI:21810956
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hc
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collectic
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LI
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution informatior
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
ate: LLCM2392  row: b  column: 19
gh quality sequence stop: 686.
       Location/Qualifiers
       1. .904
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:6249906"
       /tissue_type="ductal carcinoma, cell line"
       /lab_host="DH10B (phage-resistant)"
       /clone_lib="NIH_MGC_110"
       /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
       Site_2: EcoRI; cDNA made by oligo-dT priming.
       Directionally cloned into EcoRI/XhoI sites using the
       following 5' adaptor: GGCACGAG(G). Library constructed  by
       Ling Hong in the laboratory of Gerald M. Rubin (University
       of  California, Berkeley) using ZAP-cDNA synthesis kit
       (Stratagene) and Superscript II RT (Life Technologies).
       Note: this is a NIH_MGC Library."


            56.4%;  Score 43.4;  DB 13;  Length 904;
imilarity   72.7%;  Pred. No. 0.0043;
;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||    |   |  |    ||  |  | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 456

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 473



89619                  907 bp    mRNA    linear   EST 15-JUL-2002
NCOURT_8186629 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251590
  mRNA sequence.
89619
89619.1  GI:21814935
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 907)
-MGC http://mgc.nci.nih.gov/.
ional Institutes of Health, Mammalian Gene Collection (MGC)
ublished (1999)
tact: Robert Strausberg, Ph.D.
il: cgapbs-r@mail.nih.gov
sue Procurement: ATCC
NA Library Preparation: Rubin Laboratory
NA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
A Sequencing by: Agencourt Bioscience Corporation
one distribution: MGC clone distribution information can be
nd through the I.M.A.G.E. Consortium/LLNL at:
p://image.llnl.gov
te: LLCM2396  row: h  column: 23
gh quality sequence stop: 514.
       Location/Qualifiers
       1. .907
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:6251590"
       /tissue_type="ductal carcinoma, cell line"
       /lab_host="DH10B (phage-resistant)"
       /clone_lib="NIH_MGC_110"
       /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
       Site_2: EcoRI; cDNA made by oligo-dT priming.
       Directionally cloned into EcoRI/XhoI sites using the
       following 5' adaptor: GGCACGAG(G). Library co
       Ling Hong in the laboratory of Gerald M. Rubi
       of  California, Berkeley) using ZAP-cDNA synt
       (Stratagene) and Superscript II RT (Life Tech
       Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            56.4%;  Score 43.4;  DB 13;  Length 907;
  Best Local Similarity  72.7%;  Pred. No. 0.0043;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
              ||||||||||||||||||||||      |    | |    ||  |  | ||
Db        397 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        457 CCGTGGAGGACAAGGAT 473


RESULT 23
AW360989/c
LOCUS       AW360989                372 bp    mRNA    linear   ES
DEFINITION  PM1-CT0247-291299-007-b09 CT0247 Homo sapiens cDNA, mR
ACCESSION   AW360989
VERSION     AW360989.1  GI:6865639
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 372)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sa
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human C
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t
            291299-007-b09&t3=1999-12-29&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 244.
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0247"
                     /note="Organ: colon; Vector: puc18; Site_1: Sm
                     SmaI; A mini-library was made by cloning produ
                     from  ORESTES PCR (U.S. Letters Patent applica
                     196,716 -  Ludwig Institute for Cancer Researc
                     into  the pUC 18 vector. Reverse transcriptio
                     mRNA and cDNA amplification were performed unc
                     stringency conditions."
ORIGIN

  Query Match            54.3%;  Score 41.8;  DB 10;  Length 372;
  Best Local Similarity  71.4%;  Pred. No. 0.0094;
  Matches   55;  Conservative    0;  Mismatches   22;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||          |  |    ||  |  | |||
Db        212 AGACAATCACAGTCTCTGCGGAGCTGCACAAGCCCTCCATCTCCAGCAACAAC
```

```
CCGTGGAGGACAAGGAT 77
||||||||||||||||
CCGTGGAGGACAAGGAT 136
```

```
756277                  402 bp    mRNA    linear   EST 12-JAN-2001
3-CT0558-301000-453-b11 CT0558 Homo sapiens cDNA, mRNA sequence.
756277
756277.1  GI:12104177
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 402)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
unstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
mpson,A.J.
otgun sequencing of the human transcriptome with ORF expressed
quence tags
oc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
202663
737800
ntact: Simpson A.J.G.
boratory of Cancer Genetics
dwig Institute for Cancer Research
a Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
azil
l: +55-11-2704922
x: +55-11-2707001
ail: asimpson@ludwig.org.br
is sequence was derived from the FAPESP/LICR Human Cancer Genome
oject. This entry can be seen in the following URL
ttp://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-CT0558-
1000-453-b11&t3=2000-10-30&t4=1)
q primer: puc 18 forward
gh quality sequence stop: 402.
       Location/Qualifiers
       1. .402
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /dev_stage="Adult"
       /clone_lib="CT0558"
       /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
       SmaI; A mini-library was made by cloning products derived
       from  ORESTES PCR (U.S. Letters Patent application No.
       196,716 -  Ludwig Institute for Cancer Research) profiles
       into  the pUC 18 vector. Reverse transcription of tissue
       mRNA and cDNA amplification were performed under low
       stringency conditions."
```

```
           54.3%;  Score 41.8;  DB 10;  Length 402;
imilarity   71.4%;  Pred. No. 0.0097;
;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||| |||||||||      |    |   ||  |  | ||||||||||
AGACAATCACAGCCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 155

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 172
```

```
BM836580
LOCUS       BM836580                 444 bp    mRNA    linear   ES
DEFINITION  K-EST0112341 S9SNU601 Homo sapiens cDNA clone S9SNU601
            mRNA sequence.
ACCESSION   BM836580
VERSION     BM836580.1  GI:19192989
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S.,
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 86  row: H  column: 09
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-86-H09"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Sit
                     Site_2: XhoI; The poly (A)+ RNA was dephospho
                     bacterial alkaline phosphatase (BAP) and then
                     with tabacco acid pyrophosphatase (TAP). The
                     intact mRNA was ligated with DNA-RNA linker i
                     site by treatment of T4 RNA ligase and the fi
                     cDNA was synthesized with Superscript II usin
                     oligo-dT primer. After first strand synthesis
                     degraded by NaOH treatment and cDNA was ampli
                     reaction. The PCR products were digested with
                     cloned into DraIII- digested pME18S-FL3 vecto
                     obtained cDNA vectors were used for transform
                     competent cells E. coli Top10F' by electropor
                     The cDNA libraries constructed by this method
                     full-length enriched cDNA library."
ORIGIN
```

```
  Query Match           54.3%;  Score 41.8;  DB 12;  Length 444;
  Best Local Similarity   71.4%;  Pred. No. 0.01;
  Matches   55;  Conservative    0;  Mismatches   22;  Indels    0

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
              |||||||||||||||||||||||      |    |  |   ||  |  | ||
Db         86 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||| |||
Db        146 CCGTGGAGGACAATGAT 162


RESULT 26
AA055605
LOCUS       AA055605                 423 bp    mRNA    linear   ES
DEFINITION  zl74h11.r1 Stratagene colon (#937204) Homo sapiens cDN
            IMAGE:510405 5' similar to gb:M29540 CARCINOEMBRYONIC
```

:URSOR (HUMAN);, mRNA sequence.
;5605
;5605.1 GI:1547962

› sapiens (human)
› sapiens
ıryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ıalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 423)
lier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
ssoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
ins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
lis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
lfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
/askis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
Marra,M.
eration and analysis of 280,000 human expressed sequence tags
ome Res. 6 (9), 807-828 (1996)
44478
9549
:act: Wilson RK
ıington University School of Medicine
4 Forest Park Parkway, Box 8501, St. Louis, MO 63108
: 314 286 1800
: 314 286 1810
il: est@watson.wustl.edu
s clone is available royalty-free through LLNL ; contact the
GE Consortium (info@image.llnl.gov) for further information.
primer: -28M13 rev2 from Amersham
h quality sequence stop: 376.
Location/Qualifiers
1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3813738"
/db_xref="taxon:9606"
/clone="IMAGE:510405"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTT 3'"

54.0%; Score 41.6; DB 9; Length 423;
milarity 79.2%; Pred. No. 0.011;
Conservative 0; Mismatches 15; Indels 1; Gaps 1;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| || | || ||| || | | ||||||||||
GACAATCACAGTCTCTGCGGANTGCCCAAGCCCTCCAT-CTCCAGCAACAACTCCAAAC 102

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 119

166576 312 bp mRNA linear EST 21-APR-1997
77560 Pancreas tumor III Homo sapiens cDNA 5' end similar to
rcinoembryonic antigen (GB:M17303), mRNA sequence.
166576
166576.1 GI:2018915
.
ıo sapiens (human)
ıo sapiens
caryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 312)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Goc
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghag
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Mer
Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pell
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shi
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Colli
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Ki
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S.
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expressi
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other_ESTs: THC167592
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and express
information related to this EST, please check the TIGR
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .312
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):171263"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="Pancreas tumor III"
/note="Organ: pancreas; Vector: pBluescript SK
EcoRI; Site_2: XhoI"
ORIGIN

Query Match 53.2%; Score 41; DB 9; Length 312;
Best Local Similarity 79.2%; Pred. No. 0.016;
Matches 61; Conservative 0; Mismatches 15; Indels 1;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| || | || ||| || | | |||
Db 37 AGACAATCACAGTCTCTGCGGAGTGCCCAAGCCCTCCAT-CTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 96 CCGTGGAGGACAAGGAT 112

RESULT 28
AA132613
LOCUS AA132613 389 bp mRNA linear EST
DEFINITION zo17f02.r1 Stratagene colon (#937204) Homo sapiens cDNA
IMAGE:587163 5' similar to gb:M29540 CARCINOEMBRYONIC A
PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA132613
VERSION AA132613.1 GI:1694102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut

malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 389)
lier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
ssoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
kins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
dis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
lfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
vaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
Marra,M.
eration and analysis of 280,000 human expressed sequence tags
ome Res. 6 (9), 807-828 (1996)
44478
9549
tact: Wilson RK
hington University School of Medicine
4 Forest Park Parkway, Box 8501, St. Louis, MO 63108
: 314 286 1800
: 314 286 1810
il: est@watson.wustl.edu
s clone is available royalty-free through LLNL ; contact the
GE Consortium (info@image.llnl.gov) for further information.
ert Length: 1093 Std Error: 0.00
primer: -28M13 rev2 from Amersham
h quality sequence stop: 324.
Location/Qualifiers
1. .389
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4619465"
/db_xref="taxon:9606"
/clone="IMAGE:587163"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTT 3'"

53.2%; Score 41; DB 9; Length 389;
milarity 79.2%; Pred. No. 0.017;
Conservative 0; Mismatches 15; Indels 1; Gaps 1;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| || | || ||| || | | ||||||||||
GACAATCACAGTCTCTGCGGAGTGCCCAAGCCCTCCAT-CTCCAGCAACAACTCCAAAC 114

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 131

317244 287 bp mRNA linear EST 13-JAN-2001
2-CI0157-111200-008-g05 CI0157 Homo sapiens cDNA, mRNA sequence.
317244
817244.1 GI:12153747
T.
mo sapiens (human)
mo sapiens
karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 287)
as Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
gai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
ldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
unstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM2&t2
111200-008-g05&t3=2000-12-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 287.
FEATURES Location/Qualifiers
source 1. .287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CI0157"
/note="Organ: colon_ins; Vector: puc18; Site_1
Site_2: SmaI; A mini-library was made by cloni
derived from ORESTES PCR (U.S. Letters Patent
No. 196,716 - Ludwig Institute for Cancer Res
profiles into the pUC 18 vector. Reverse tran
tissue mRNA and cDNA amplification were perfor
low stringency conditions."
ORIGIN

Query Match 52.2%; Score 40.2; DB 10; Length 287;
Best Local Similarity 70.1%; Pred. No. 0.027;
Matches 54; Conservative 0; Mismatches 23; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | | | |||
Db 279 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCGCCAACAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 219 CCGTGGAGGACAAGGAT 203

RESULT 30
AA134992
LOCUS AA134992 327 bp mRNA linear EST
DEFINITION zo23a06.r1 Stratagene colon (#937204) Homo sapiens cDN
IMAGE:587698 5' similar to gb:M29540 CARCINOEMBRYONIC
PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA134992
VERSION AA134992.1 GI:1696111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H
REFERENCE 1 (bases 1 to 327)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiape
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thi
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R.
and Marra,M.
TITLE Generation and analysis of 280,000 human expressed seq
JOURNAL Genome Res. 6 (9), 807-828 (1996)

4478
549
act: Wilson RK
ington University School of Medicine
Forest Park Parkway, Box 8501, St. Louis, MO 63108
314 286 1800
314 286 1810
l: est@watson.wustl.edu
clone is available royalty-free through LLNL ; contact the
E Consortium (info@image.llnl.gov) for further information.
e considered overall poor quality
rt Length: 1162 Std Error: 0.00
primer: -28M13 rev2 from Amersham
quality sequence stop: 1.
Location/Qualifiers
1. .327
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4620000"
/db_xref="taxon:9606"
/clone="IMAGE:587698"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTT 3'"

52.2%; Score 40.2; DB 9; Length 327;
ilarity 76.6%; Pred. No. 0.028;
Conservative 0; Mismatches 17; Indels 1; Gaps 1;

:ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||| |||||||||||||||| || | || ||| || | | | ||||||||
:ACANTCACAGTCTCTGCGGAGTGCCCAAGCNCTCCAT-CTCCAGCNACANCTCCAAAC 114

:GTGGAGGACAAGGAT 77
|||||||||||||||
:GTGGAGGACAAGGAT 131

)0185 630 bp mRNA linear EST 04-SEP-2003
:p686F02218_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
:p686F02218 5', mRNA sequence.
)0185
)0185.1 GI:31916654

> sapiens (human)
> sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 630)
ecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
>,G., Han,M. and Wiemann,S.
(Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
ublished (2003)
tact: MIPS
3
olstaedter Landstr.1, D-85764 Neuherberg, Germany
s is the 5' sequence of the clone insert
ne from S. Wiemann, Molecular Genome Analysis, German Cancer
earch Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
uenced by GBF (National Research Centre for Biotechnology Ltd.,
unschweig/Germany) within the cDNA sequencing consortium of the
man Genome Project.
s1 sequence available.

This clone (DKFZp686F02218) is available at the RZPD in
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES Location/Qualifiers
source 1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686F02218"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2
cDNA-collection"
ORIGIN

Query Match 52.2%; Score 40.2; DB 13; Length 630;
Best Local Similarity 70.1%; Pred. No. 0.038;
Matches 54; Conservative 0; Mismatches 23; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||| |||||||||||||||| | | | || | | ||||
Db 86 AGACGGTCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 146 CCGTGGAGGACAAGGAT 162

RESULT 32
BF084626
LOCUS BF084626 494 bp mRNA linear EST
DEFINITION RC1-CT0524-150900-022-f05 CT0524 Homo sapiens cDNA, mRN
ACCESSION BF084626
VERSION BF084626.1 GI:10878469
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor
REFERENCE 1 (bases 1 to 494)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., B
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Cost
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., S
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de So
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC
900-022-f05&t3=2000-09-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 9
High quality sequence stop: 494.
FEATURES Location/Qualifiers
source 1. .494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"

/clone_lib="CT0524"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

49.9%; Score 38.4; DB 10; Length 494;
nilarity 70.8%; Pred. No. 0.13;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

TCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACCCGTG 65
||||||||||||||||| | | || | | |||||||||||||||
TCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAACCCGTG 66

AGGACAAGGAT 77
||||||||||
AGGACAAGGAT 78

52017 623 bp mRNA linear EST 19-MAY-2000
-CT0347-160200-013-h04 CT0347 Homo sapiens cDNA, mRNA sequence.
52017
52017.1 GI:7957710
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 623)
s Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
ai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
dman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
nstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
are,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
pson,A.J.
tgun sequencing of the human transcriptome with ORF expressed
uence tags
c. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
tact: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-CT0347-160
-013-h04&t3=2000-02-16&t4=1)
primer: puc 18 forward
h quality sequence start: 22
h quality sequence stop: 623.
Location/Qualifiers
1. .623
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0347"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed unde stringency conditions."

ORIGIN

Query Match 49.9%; Score 38.4; DB 10; Length 623;
Best Local Similarity 70.8%; Pred. No. 0.14;
Matches 51; Conservative 0; Mismatches 21; Indels 0;

Qy 6 ATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAA
||||||||||||||||| | | || | | ||||||||
Db 104 ATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAA

Qy 66 GAGGACAAGGAT 77
||||||||||||
Db 164 GAGGACAAGGAT 175

RESULT 34
AA127122
LOCUS AA127122 440 bp mRNA linear EST
DEFINITION zl87f02.r1 Stratagene colon (#937204) Homo sapiens cDNA
IMAGE:511611 5' similar to gb:M29540 CARCINOEMBRYONIC A
PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA127122
VERSION AA127122.1 GI:1686499
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1 (bases 1 to 440)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapel
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thie
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R.,
and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequ
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; con
IMAGE Consortium (info@image.llnl.gov) for further info
Insert Length: 1245 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 361.
FEATURES Location/Qualifiers
source 1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3844106"
/db_xref="taxon:9606"
/clone="IMAGE:511611"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene colon (#937204)"
/note="Organ: colon; Vector: pBluescript SK-; EcoRI; Site_2: XhoI; Cloned unidirectionally. Oligo dT. T-84 colonic epithelial cell line. insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' ad sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor seq CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

Query Match 48.8%; Score 37.6; DB 9; Length 440;

```
ilarity   75.3%;  Pred. No. 0.21;
 Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

:ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
 |||||||||||||||||||      ||  | || |||     |  | ||||||||||
:ACAATCACAGTCTCTGCGGNGTGCCCAAGCCCTCCAT--NNCAGCAACAACTCCAAAC 277

:GTGGAGGACAAGGAT 77
 |||||||||||||||
:GTGGAGGACAAGGAT 294
```

```
:7770                 541 bp    mRNA    linear   EST 30-MAY-2000
DT0045-140200-082-g08 DT0045 Homo sapiens cDNA, mRNA sequence.
7770
7770.1  GI:8113198

 sapiens (human)
 sapiens
.ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ialia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
bases 1 to 541)
 Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
i,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
lman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
istein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
ire,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
)son,A.J.
:gun sequencing of the human transcriptome with ORF expressed
ience tags
:. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
)2663
:7800
:act: Simpson A.J.G.
)ratory of Cancer Genetics
/ig Institute for Cancer Research
 Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
:il
  +55-11-2704922
  +55-11-2707001
.l: asimpson@ludwig.org.br
 sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
:p://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-DT0045-140
-082-g08&t3=2000-02-14&t4=1)
 primer: puc 18 forward
 quality sequence start: 20
 quality sequence stop: 75.
     Location/Qualifiers
     1. .541
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /dev_stage="Adult"
     /clone_lib="DT0045"
     /note="Organ: denis_drash; Vector: puc18; Site_1: SmaI;
     Site_2: SmaI; A mini-library was made by cloning products
     derived from  ORESTES PCR (U.S. Letters Patent application
     No. 196,716 -  Ludwig Institute for Cancer Research)
     profiles into  the pUC 18 vector. Reverse transcription of
     tissue mRNA and cDNA amplification were performed under
     low stringency conditions."


         48.1%;  Score 37;  DB 10;  Length 541;
milarity   67.5%;  Pred. No. 0.36;
  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||      |  |  |  |  | |||||||| |
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCACCAGCAACAACTCCAACC 428

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||| |||||
Db        429 CCGTGGAGGATGAGGAT 445


RESULT 36
AW361304
LOCUS       AW361304                 590 bp    mRNA    linear   EST
DEFINITION  RC3-CT0254-280999-011-a08 CT0254 Homo sapiens cDNA, mRN
ACCESSION   AW361304
VERSION     AW361304.1  GI:6865954
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor
REFERENCE   1  (bases 1 to 590)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Ca
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC3&t2=
            280999-011-a08&t3=1999-09-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0254"
                     /note="Organ: colon; Vector: puc18; Site_1: Sm
                     SmaI; A mini-library was made by cloning produ
                     from  ORESTES PCR (U.S. Letters Patent applica
                     196,716 -  Ludwig Institute for Cancer Researc
                     into  the pUC 18 vector. Reverse transcription
                     mRNA and cDNA amplification were performed und
                     stringency conditions."
ORIGIN

  Query Match           45.7%;  Score 35.2;  DB 10;  Length 590;
  Best Local Similarity   73.8%;  Pred. No. 1.4;
  Matches   59;  Conservative    0;  Mismatches   18;  Indels    3;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGT---CAGGTCAT
              |||| |||||||||||||||  |  |    | |   ||   ||||  |
Db         10 AGACGATCACAGTCTCTGCGGAGCCGCCCAAGCCCGTCCAGTCTCCAGGCAAC

Qy         58 AACCCGTGGAGGACAAGGAT 77
              ||||||||||||||||||||
Db         70 AACCCGTGGAGGACAAGGAT 89


RESULT 37
BG950148/c
LOCUS       BG950148                 296 bp    mRNA    linear   EST
DEFINITION  CM2-CT0639-030101-681-e08 CT0639 Homo sapiens cDNA, mRN
ACCESSION   BG950148
VERSION     BG950148.1  GI:14368319
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 296)
s Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
ii,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
lman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
istein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
are,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
son,A.J.
gun sequencing of the human transcriptome with ORF expressed
ience tags
. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
02663
37800
act: Simpson A.J.G.
oratory of Cancer Genetics
wig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
zil
: +55-11-2704922
: +55-11-2707001
il: asimpson@ludwig.org.br
s sequence was derived from the FAPESP/LICR Human Cancer Genome
ject. This entry can be seen in the following URL
tp://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-CT0639-
.01-681-e08&t3=2001-01-03&t4=1)
primer: puc 18 forward
n quality sequence start: 29
n quality sequence stop: 293.
Location/Qualifiers
1. .296
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0639"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

45.5%; Score 35; DB 12; Length 296;
nilarity 66.7%; Pred. No. 1.2;
Conservative 0; Mismatches 25; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||| |||||||||| |||||| | | | || | | ||||||||||
GACGATCACAGTCTATGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 106

CGTGGAGGACAAGG 75
|||||||||| | |
CGTGGAGGACCAAG 91

35665 914 bp mRNA linear EST 15-SEP-2000
304416F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3639027 5',
A sequence.
35665
35665.1 GI:10149657
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 914)
-MGC http://mgc.nci.nih.gov/.
ional Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LL
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM340 row: p column: 04
High quality sequence stop: 705.
FEATURES Location/Qualifiers
source 1. .914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3639027"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/note="Organ: pancreas; Vector: pOTB7; Site_1:
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites usi
following 5' adaptor: GGCACGAG(G). Library co
by Ling Hong in the laboratory of Gerald M. Ru
(University of California, Berkeley) using ZA
synthesis kit (Stratagene) and Superscript II
Technologies)."
ORIGIN

Query Match 44.9%; Score 34.6; DB 10; Length 914;
Best Local Similarity 74.0%; Pred. No. 2.6;
Matches 57; Conservative 0; Mismatches 19; Indels 1;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||||||||||||||||||||| | | || | ||| |||
Db 521 AGACAATCACAGTCTCTGCGGAGCTGACCAAGCCCTCCATCTCCAG-CAAAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 580 CCGTGGAGGACAAGGAT 596

RESULT 39
AA327327
LOCUS AA327327 237 bp mRNA linear EST
DEFINITION EST30641 Colon I Homo sapiens cDNA 5' end similar to
carcinoembryonic antigen (GB:M17303), mRNA sequence.
ACCESSION AA327327
VERSION AA327327.1 GI:1979572
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor
REFERENCE 1 (bases 1 to 237)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Goc
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghag
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Mer
Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pell
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shi
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Colli
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kir
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A.,

ser,C.M. and Venter,J.C.
tial assessment of human gene diversity and expression patterns
ed upon 83 million nucleotides of cDNA sequence
ure 377 (6547 Suppl), 3-174 (1995)
26280
6098
er_ESTs: THC167592
tact: Kerlavage, AR
informatics
Institute for Genomic Research
2 Medical Center Drive, Rockville, MD 20850 USA
: 3018699056
: 3018699423
il: arkerlav@tigr.org
clone availability, additional sequence and expression
ormation related to this EST, please check the TIGR Human Gene
ex (http://www.tigr.org/tdb/hgi/hgi.html)
primer: M13 Reverse.
Location/Qualifiers
1. .237
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):128005"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="Colon I"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

44.7%; Score 34.4; DB 9; Length 237;
ilarity 65.8%; Pred. No. 1.7;
Conservative 0; Mismatches 26; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATGACCAGCAACAACTCCAACCC 221

TGGAGGACAAGGAT 77
TGGAGGATGAGGAT 237

9404 423 bp mRNA linear EST 19-MAY-2000
CT0215-160300-089-F08 CT0215 Homo sapiens cDNA, mRNA sequence.
9404
9404.1 GI:7944921

sapiens (human)
sapiens
ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
bases 1 to 423)
Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
i,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
man,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
stein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
re,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
son,A.J.
gun sequencing of the human transcriptome with ORF expressed
ence tags
. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2663
7800
act: Simpson A.J.G.
ratory of Cancer Genetics
ig Institute for Cancer Research
Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
il
+55-11-2704922
+55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Ca
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL
300-089-F08&t3=2000-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 362.

FEATURES Location/Qualifiers
source 1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0215"
/note="Organ: colon; Vector: puc18; Site_1: Sm
SmaI; A mini-library was made by cloning produ
from ORESTES PCR (U.S. Letters Patent applica
196,716 - Ludwig Institute for Cancer Researc
into the pUC 18 vector. Reverse transcription
mRNA and cDNA amplification were performed und
stringency conditions."

ORIGIN

Query Match 44.7%; Score 34.4; DB 10; Length 423;
Best Local Similarity 65.8%; Pred. No. 2.1;
Matches 50; Conservative 0; Mismatches 26; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
Db 228 GACGATCACAGTCTATGCAGAGCCACCCAGACCCTTCATCACCAGCAACAACT

Qy 62 CGTGGAGGACAAGGAT 77
Db 288 CGTGGAGGATAAGGAT 303

Search completed: April 13, 2004, 20:14:19
Job time : 1858.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

leic search, using sw model

April 13, 2004, 19:12:18 ; Search time 174.5 Seconds
(without alignments)
1655.373 Million cell updates/sec

US-10-090-326-24
77
1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

2475585 seqs, 1875730760 residues

hits satisfying chosen parameters: 4951170

ength: 0
ength: 2000000000

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

is the number of results predicted by chance to have a
ter than or equal to the score of the result being printed,
ived by analysis of the total score distribution.

SUMMARIES

	% Query Match	Length	DB	ID	Description
	58.7	1943	14	US-10-106-698-594	Sequence 594, App
	58.4	1623	13	US-10-001-873-14	Sequence 14, Appl
	57.9	2459	9	US-09-925-301-291	Sequence 291, App
	56.4	111	9	US-09-815-343-268	Sequence 268, App
	56.4	111	12	US-10-097-105-268	Sequence 268, App
	56.4	190	9	US-09-815-343-254	Sequence 254, App
	56.4	190	12	US-10-097-105-254	Sequence 254, App
	56.4	326	9	US-09-815-343-887	Sequence 887, App
	56.4	326	9	US-09-815-343-1112	Sequence 1112, Ap
	56.4	326	9	US-09-815-343-1390	Sequence 1390, Ap
	56.4	326	12	US-10-097-105-887	Sequence 887, App
	56.4	326	12	US-10-097-105-1112	Sequence 1112, Ap
	56.4	326	12	US-10-097-105-1390	Sequence 1390, Ap
	56.4	532	9	US-09-998-598-780	Sequence 780, App
	56.4	571	9	US-09-815-343-725	Sequence 725, App
	56.4	571	12	US-10-097-105-725	Sequenc
	56.4	572	9	US-09-815-343-1124	Sequenc
	56.4	572	12	US-10-097-105-1124	Sequenc
	56.4	2106	15	US-10-168-417A-3	Sequenc
	56.4	2220	9	US-09-756-551A-16	Sequenc
	56.4	2349	14	US-10-267-384-145	Sequenc
	56.4	2364	14	US-10-198-846-10292	Sequenc
	56.4	2434	14	US-10-267-384-144	Sequenc
	56.4	2728	10	US-09-814-353-20888	Sequenc
	56.4	2974	9	US-09-954-456-56	Sequenc
	56.4	2974	9	US-09-880-107-2317	Sequenc
	56.4	2974	12	US-10-240-425-1314	Sequenc
	56.4	2974	14	US-10-157-031-340	Sequenc
	56.4	2974	14	US-10-207-655-86	Sequenc
	56.4	2974	15	US-10-117-937-593	Sequenc
	56.4	2974	15	US-10-295-027-331	Sequenc
	53.2	588	10	US-09-871-161-447	Sequenc
	50.6	710	14	US-10-066-543-95	Sequenc
c	42.6	456	9	US-09-998-598-1691	Sequence
c	42.6	472	9	US-09-736-457-1050	Sequence
c	42.6	472	9	US-09-902-941-1050	Sequence
c	42.6	472	9	US-09-849-626-1050	Sequence
c	42.6	472	14	US-10-017-754-1050	Sequenc
c	42.6	472	14	US-10-113-872-1050	Sequenc
c	42.6	476	9	US-09-920-300A-435	Sequence
c	42.6	476	9	US-09-998-598-234	Sequence
c	42.6	476	13	US-10-033-528-435	Sequenc
c	42.6	476	14	US-10-099-926-435	Sequenc
	42.6	478	9	US-09-920-300A-678	Sequenc
c	42.6	478	9	US-09-998-598-293	Sequenc

ALIGNMENTS

RESULT 1
US-10-106-698-594
; Sequence 594, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleot
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 594
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1177)..(1177)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-594

Query Match 58.7%; Score 45.2; DB 14; Length 1943;
Best Local Similarity 72.7%; Pred. No. 4.7e-07;
Matches 56; Conservative 1; Mismatches 20; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              |||||||||||||||||||| :  |    | |   ||  |  | ||||
Db       1155 AGACAATCACAGTCTCTGCGGANSTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
```

'CGTGGAGGACAAGGAT 1231

4
Application US/10001873
o. US20020160388A1
MATION:
acina, Roberto
Recipon, Herve
Chen, Sei-Yu
Sun, Yongming
Liu, Chenghua
Turner, Leah
ENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
CE: DEX-0275
ICATION NUMBER: US/10/001,873
NG DATE: 2001-11-20
ATION NUMBER: 60/252,055
DATE: 2000-11-20
ATION NUMBER: 60/252,496
DATE: 2000-11-22
Q ID NOS: 55
tentIn version 3.1

3

omo sapien

isc_feature
856)..(856)
MATION: a, c, g or t
4

58.4%; Score 45; DB 13; Length 1623;
nilarity 74.0%; Pred. No. 5.3e-07;
Conservative 0; Mismatches 20; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | | || | | |||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 110

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 127

91
Application US/09925301
20020052308A1
MATION:
osen et al.
ENTION: Nucleic Acids, Proteins and Antibodies
CE: PA106
ICATION NUMBER: US/09/925,301
NG DATE: 2001-08-10
ATION NUMBER: PCT/US00/05882
DATE: 2000-03-08
ATION NUMBER: 60/124,270
DATE: 1999-03-12
Q ID NOS: 1694
tentIn Ver. 2.0

9

omo sapiens

isc_feature
3)
MATION: n equals a,t,g, or c
isc_feature
1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1604)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1605)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2374)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2392)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-291

Query Match 57.9%; Score 44.6; DB 9; Length 2459;
Best Local Similarity 72.7%; Pred. No. 8.4e-07;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 1582 AGACAATCACAGTCTCTGCGGANNTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1642 CCGTGGAGGACAAGGAT 1658

RESULT 4
US-09-815-343-268
; Sequence 268, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 111
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-815-343-268

Query Match 56.4%; Score 43.4; DB 9; Length 111;
Best Local Similarity 72.7%; Pred. No. 1.1e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 5
US-10-097-105-268
; Sequence 268, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

'ENTION: AND DIAGNOSIS OF COLON CANCER
ICE: 210121.504C1
.ICATION NUMBER: US/10/097,105
NG DATE: 2002-03-13
Q ID NOS: 1562
stSEQ for Windows Version 4.0

omo sapiens
68

56.4%; Score 43.4; DB 12; Length 111;
milarity 72.7%; Pred. No. 1.1e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97

54
Application US/09815343
20010055596A1
MATION:
eagher, Madeleine
Xu, Jiangchun
King, Gordon E.
ENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
ENTION: DIAGNOSIS OF COLON CANCER
CE: 210121.504
ICATION NUMBER: US/09/815,343
NG DATE: 2001-03-22
Q ID NOS: 1556
stSEQ for Windows Version 4.0

omo sapien
54

56.4%; Score 43.4; DB 9; Length 190;
ilarity 72.7%; Pred. No. 1.2e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 97

4
Application US/10097105
. US20040037842A1
MATION:
agher, Madeleine Joy
ing, Gordon E.
ecrist, Heather
arlocker, Susan L.
u, Jiangchun
NTION: COMPOSITIONS AND METHODS FOR THE THERAPY
NTION: AND DIAGNOSIS OF COLON CANCER
E: 210121.504C1
CATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 254
; LENGTH: 190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-105-254

Query Match 56.4%; Score 43.4; DB 12; Length 190;
Best Local Similarity 72.7%; Pred. No. 1.2e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | || | | |||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 8
US-09-815-343-887
; Sequence 887, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 887
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-815-343-887

Query Match 56.4%; Score 43.4; DB 9; Length 326;
Best Local Similarity 72.7%; Pred. No. 1.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | || | | |||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 9
US-09-815-343-1112
; Sequence 1112, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1112
; LENGTH: 326

lomo sapien
.112

56.4%; Score 43.4; DB 9; Length 326;
.milarity 72.7%; Pred. No. 1.4e-06;
: Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||| | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 97

.390
), Application US/09815343
320010055596A1
MATION:
leagher, Madeleine
Xu, Jiangchun
King, Gordon E.
ENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
ENTION: DIAGNOSIS OF COLON CANCER
CE: 210121.504
ICATION NUMBER: US/09/815,343
NG DATE: 2001-03-22
Q ID NOS: 1556
stSEQ for Windows Version 4.0
0

omo sapien
390

56.4%; Score 43.4; DB 9; Length 326;
milarity 72.7%; Pred. No. 1.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||| | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||||
CGTGGAGGACAAGGAT 97

87
Application US/10097105
o. US20040037842A1
MATION:
eagher, Madeleine Joy
King, Gordon E.
Secrist, Heather
Harlocker, Susan L.
Xu, Jiangchun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF COLON CANCER
CE: 210121.504C1
ICATION NUMBER: US/10/097,105
NG DATE: 2002-03-13
Q ID NOS: 1562
stSEQ for Windows Version 4.0

omo sapiens
37

Query Match 56.4%; Score 43.4; DB 12; Length 326;
Best Local Similarity 72.7%; Pred. No. 1.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| | | || | | ||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 12
US-10-097-105-1112
; Sequence 1112, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1112
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-105-1112

Query Match 56.4%; Score 43.4; DB 12; Length 326;
Best Local Similarity 72.7%; Pred. No. 1.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | || | | ||||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 13
US-10-097-105-1390
; Sequence 1390, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1390
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-105-1390

56.4%; Score 43.4; DB 12; Length 326;
nilarity 72.7%; Pred. No. 1.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97

30
Application US/09998598
20020150922A1
MATION:
olk, John A.
Xu, Jiangchun
Chenault, Ruth A.
Meagher, Madelein Joy
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF COLON CANCER
E: 210121.561
ICATION NUMBER: US/09/998,598
NG DATE: 2001-11-16
) ID NOS: 2606
rixa Invention Disclosure Database

mo sapiens
10

56.4%; Score 43.4; DB 9; Length 532;
ilarity 72.7%; Pred. No. 1.6e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 97

5
Application US/09815343
0010055596A1
ATION:
agher, Madeleine
u, Jiangchun
ing, Gordon E.
NTION: COMPOSITIONS AND METHODS FOR THERAPY AND
NTION: DIAGNOSIS OF COLON CANCER
E: 210121.504
CATION NUMBER: US/09/815,343
G DATE: 2001-03-22
ID NOS: 1556
tSEQ for Windows Version 4.0

mo sapien

sc_feature
)...(571)
ATION: n = A,T,C or G
5

Query Match 56.4%; Score 43.4; DB 9; Length 571;
Best Local Similarity 72.7%; Pred. No. 1.7e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 16
US-10-097-105-725
; Sequence 725, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 725
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 483, 520, 523, 535, 538, 539, 549, 558, 563, 566
; OTHER INFORMATION: n = A,T,C or G
US-10-097-105-725

Query Match 56.4%; Score 43.4; DB 12; Length 571;
Best Local Similarity 72.7%; Pred. No. 1.7e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 17
US-09-815-343-1124
; Sequence 1124, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1124
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:

aisc_feature
(1)...(572)
RMATION: n = A,T,C or G
1124

56.4%; Score 43.4; DB 9; Length 572;
imilarity 72.7%; Pred. No. 1.7e-06;
: Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

:CGTGGAGGACAAGGAT 77
:CGTGGAGGACAAGGAT 97

124
, Application US/10097105
o. US20040037842A1
MATION:
eagher, Madeleine Joy
King, Gordon E.
Secrist, Heather
Harlocker, Susan L.
Xu, Jiangchun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF COLON CANCER
CE: 210121.504C1
ICATION NUMBER: US/10/097,105
NG DATE: 2002-03-13
Q ID NOS: 1562
stSEQ for Windows Version 4.0
4

omo sapiens

isc_feature
58, 313, 355, 378, 414, 419, 432, 456, 458, 468, 478, 483,
90, 505, 522, 526, 530, 537, 539, 540, 548, 561, 562, 567,
71
MATION: n = A,T,C or G
124

56.4%; Score 43.4; DB 12; Length 572;
ilarity 72.7%; Pred. No. 1.7e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

GTGGAGGACAAGGAT 77
GTGGAGGACAAGGAT 97

plication US/10168417A
. US20040009185A1
MATION:
tage, Peter
arber, Brian
ambhara, Suryprakash
ia, Charles Dwo Yuan
NTION: Enhancing the Immune Response to an Antigen by Presensitizing wit
NTION: Inducing Agent Prior to Immunizing with the Inducing Agent and t
E: 11014-18-US
CATION NUMBER: US/10/168,417A

; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/174,587
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: PCT/CA01/00005
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: modified CEA
US-10-168-417A-3

Query Match 56.4%; Score 43.4; DB 15; Length 2106;
Best Local Similarity 72.7%; Pred. No. 2.3e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
Db 1472 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1532 CCGTGGAGGACAAGGAT 1548

RESULT 20
US-09-756-551A-16
; Sequence 16, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:

IE: (617) 227-7400
(617) 742-4214
FOR SEQ ID NO: 16:
HARACTERISTICS:
2220 base pairs
ucleic acid
)NESS: single
': linear
'YPE: cDNA

': CDS
I: 1..2203
16

56.4%; Score 43.4; DB 9; Length 2220;
milarity 72.7%; Pred. No. 2.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1630

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1647

45
Application US/10267384
). US20030198623A1
)RMATION:
ANT: Paoletti, Enzo
Tartaglia, James
Cox, William I.
)F INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
OF SEQUENCES: 217
PONDENCE ADDRESS:
)DRESSEE: Curtis, Morris & Safford
TREET: 530 Fifth Avenue
ITY: New York
TATE: NY
)UNTRY: USA
IP: 10036
ER READABLE FORM:
)DIUM TYPE: Floppy disk
)MPUTER: IBM PC compatible
)ERATING SYSTEM: PC-DOS/MS-DOS
)FTWARE: PatentIn Release #1.0, Version #1.25
APPLICATION DATA:
PLICATION NUMBER: US/10/267,384
LING DATE: 09-Oct-2002
ASSIFICATION: 435
Y/AGENT INFORMATION:
ME: Frommer, William S.
GISTRATION NUMBER: 25,506
FERENCE/DOCKET NUMBER: 454310-2530
MUNICATION INFORMATION:
LEPHONE: (212) 840-3333
LEFAX: (212) 840-0712
LEX: 425066CURTMS
FOR SEQ ID NO: 145:
E CHARACTERISTICS:
NGTH: 2349 base pairs
PE: nucleic acid
RANDEDNESS: single
POLOGY: linear
E TYPE: cDNA
E DESCRIPTION: SEQ ID NO: 145:
5

56.4%; Score 43.4; DB 14; Length 2349;
ilarity 72.7%; Pred. No. 2.4e-06;

Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||| | | | || | | |||
Db 1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1718 CCGTGGAGGACAAGGAT 1734

RESULT 22
US-10-198-846-10292
; Sequence 10292, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION,
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10292
; LENGTH: 2364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-10292

Query Match 56.4%; Score 43.4; DB 14; Length 2364;
Best Local Similarity 72.7%; Pred. No. 2.4e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||| | | | || | | |||
Db 1717 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1777 CCGTGGAGGACAAGGAT 1793

RESULT 23
US-10-267-384-144
; Sequence 144, Application US/10267384
; Publication No. US20030198623A1
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; Tartaglia, James
; Cox, William I.
; TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
; NUMBER OF SEQUENCES: 217
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/267,384

FILING DATE: 09-Oct-2002
CLASSIFICATION: 435
NEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
OMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CURTMS
N FOR SEQ ID NO: 144:
NCE CHARACTERISTICS:
LENGTH: 2434 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
OLE TYPE: cDNA
NCE DESCRIPTION: SEQ ID NO: 144:
144

56.4%; Score 43.4; DB 14; Length 2434;
milarity 72.7%; Pred. No. 2.4e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1714

CCGTGGAGGACAAGGAT 77
CCGTGGAGGACAAGGAT 1731

0888
8, Application US/09814353
o. US20030165831A1
MATION:
ee, John
Thompson, Pamela
Lillie, James
ENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
ENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
ENTION: THERAPY OF OVARIAN CANCER
CE: MRI-006B
ICATION NUMBER: US/09/814,353
NG DATE: 2001-03-21
ATION NUMBER: US 60/191,031
DATE: 2000-03-21
ATION NUMBER: US 60/207,124
DATE: 2000-05-25
ATION NUMBER: US 60/211,940
DATE: 2000-06-15
ATION NUMBER: US 60/216,820
DATE: 2000-07-07
ATION NUMBER: US 60/220,661
DATE: 2000-07-25
ATION NUMBER: US 60/257,672
DATE: 2000-12-21
Q ID NOS: 22037
stSEQ for Windows Version 4.0
88
8

omo sapiens
0888

56.4%; Score 43.4; DB 10; Length 2728;
milarity 72.7%; Pred. No. 2.5e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60

Db 1753 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1813 CCGTGGAGGACAAGGAT 1829

RESULT 25
US-09-954-456-56
; Sequence 56, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therape
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-56

Query Match 56.4%; Score 43.4; DB 9; Length 2974;
Best Local Similarity 72.7%; Pred. No. 2.5e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC

Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 26
US-09-880-107-2317
; Sequence 2317, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379

; DATE: 2000-06-14
:ATION NUMBER: US 60/237,054
; DATE: 2000-10-02
:Q ID NOS: 3950
itentIn Ver. 2.1
.7
'4

Homo sapiens

MATION: Genbank Accession No. US20020142981A1 M29540
317

56.4%; Score 43.4; DB 9; Length 2974;
milarity 72.7%; Pred. No. 2.5e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1665

314
. Application US/10240425
). US20040033502A1
4ATION:
illiams, Amanda
Boland, Joseph F.
Lord, Reginald V.
Alvarez, Chris
Metzel, Jon C.
Scherf, Uwe
Vockley, Joseph G.
ENTION: Gene Expression Profiles in Esophageal Tissue
E: 44921-5026
CATION NUMBER: US/10/240,425
IG DATE: 2002-09-30
ATION NUMBER: PCT/US01/09847
DATE: 2001-03-28
ATION NUMBER: US 60/193,446
DATE: 2000-03-31
ID NOS: 1588
entIn Ver. 2.1

mo sapiens

ATION: Genbank Accession No. US20040033502A1 M29540
14

56.4%; Score 43.4; DB 12; Length 2974;
ilarity 72.7%; Pred. No. 2.5e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | || | | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1665

0
Application US/10157031
. US20030108890A1

; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associate
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-157-031-340

Query Match 56.4%; Score 43.4; DB 14; Length 2974;
Best Local Similarity 72.7%; Pred. No. 2.5e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||||||||||||||||||||| | | || | | |||
Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 29
US-10-207-655-86
; Sequence 86, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEIN
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 86
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-86

Query Match 56.4%; Score 43.4; DB 14; Length 2974;
Best Local Similarity 72.7%; Pred. No. 2.5e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
||||||||||||||||||||| | | || | | ||||
Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 30
US-10-117-937-593
; Sequence 593, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong

VENTION: EPITOPE SEQUENCES
NCE: CTLIMM.027A
LICATION NUMBER: US/10/117,937
ING DATE: 2002-04-04
CATION NUMBER: US 60/282,211
3 DATE: 2001-04-06
CATION NUMBER: US 60/337,017
3 DATE: 2001-11-07
CATION NUMBER: US 60/363,210
3 DATE: 2002-03-07
EQ ID NOS: 602
astSEQ for Windows Version 4.0
3
74

Homo sapiens
593

56.4%; Score 43.4; DB 15; Length 2974;
imilarity 72.7%; Pred. No. 2.5e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

CCGTGGAGGACAAGGAT 77
CCGTGGAGGACAAGGAT 1665

31
Application US/10295027
No. US20030232350A1
MATION:
Afar, Daniel
Aziz, Natasha
Ginsberg, Wendy M.
Gish, Kurt C.
Glynne, Richard
Hevezi, Peter A.
Mack, David H.
Murray, Richard
Watson, Susan R.
Eos Biotechnology, Inc.
ENTION: Methods of Diagnosis of Cancer, Compositions and
ENTION: Methods of Screening for Modulators of Cancer
CE: 018501-012500US
ICATION NUMBER: US/10/295,027
NG DATE: 2002-11-13
ATION NUMBER: US 09/663,733
DATE: 2000-09-15
ATION NUMBER: US 60/350,666
DATE: 2001-11-13
ATION NUMBER: US 60/335,394
DATE: 2001-11-15
ATION NUMBER: US 60/332,464
DATE: 2001-11-21
ATION NUMBER: US 60/334,393
DATE: 2001-11-29
ATION NUMBER: US 60/340,376
DATE: 2001-12-14
ATION NUMBER: US 60/347,211
DATE: 2002-01-08
ATION NUMBER: US 60/347,349
DATE: 2002-01-10
ATION NUMBER: US 60/355,250
DATE: 2002-02-08
ATION NUMBER: US 60/356,714
DATE: 2002-02-13
ior Application data removed - See File Wrapper or PALM.
Q ID NOS: 1386

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 331
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-331

Query Match 56.4%; Score 43.4; DB 15; Length 2974
Best Local Similarity 72.7%; Pred. No. 2.5e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 32
US-09-871-161-447
; Sequence 447, Application US/09871161
; Publication No. US20030097666A1
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/871,161
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/328,111
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/117,393
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 60/098,639
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 447
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(588)
; OTHER INFORMATION: n = A,T,C or G
US-09-871-161-447

Query Match 53.2%; Score 41; DB 10; Length 588;
Best Local Similarity 79.2%; Pred. No. 1.4e-05;
Matches 61; Conservative 0; Mismatches 15; Indels 1;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
Db 314 AGACAATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCAT-CTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
Db 373 CCGTGGAGGACAAGGAT 389

RESULT 33
US-10-066-543-95
; Sequence 95, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather

Carter, Darrick
Fanger, Gary R.
Smith, Carole L.
Durham, Margarita
Stolk, John A.
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF COLON CANCER
CE: 210121.563
ICATION NUMBER: US/10/066,543
NG DATE: 2002-01-31
Q ID NOS: 3417
stSEQ for Windows Version 4.0

omo sapiens

isc_feature
, 2, 551, 552, 616, 642, 647, 683, 699
MATION: n = A,T,C or G
5

50.6%; Score 39; DB 14; Length 710;
milarity 76.6%; Pred. No. 8.4e-05;
Conservative 0; Mismatches 17; Indels 1; Gaps 1;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
GACAATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCAT-CTCCAGCAACNNCTCCAAAC 560

CGTGGAGGACAAGGAT 77
CGTGGAGGACAAGGAT 577

591/c
, Application US/09998598
20020150922A1
MATION:
tolk, John A.
Ku, Jiangchun
Chenault, Ruth A.
Meagher, Madelein Joy
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF COLON CANCER
CE: 210121.561
ICATION NUMBER: US/09/998,598
NG DATE: 2001-11-16
Q ID NOS: 2606
rixa Invention Disclosure Database
L

omo sapiens
591

42.6%; Score 32.8; DB 9; Length 456;
ilarity 64.5%; Pred. No. 0.018;
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
GTGGAGGATGAGGAT 94

)50/c
Application US/09736457

; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AN
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1050
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-1050

Query Match 42.6%; Score 32.8; DB 9; Length 472;
Best Local Similarity 64.5%; Pred. No. 0.018;
Matches 49; Conservative 0; Mismatches 27; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
Db 169 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy 62 CGTGGAGGACAAGGAT 77
Db 109 CGTGGAGGATGAGGAT 94

RESULT 36
US-09-902-941-1050/c
; Sequence 1050, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1050
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1050

Query Match 42.6%; Score 32.8; DB 9; Length 472;
Best Local Similarity 64.5%; Pred. No. 0.018;
Matches 49; Conservative 0; Mismatches 27; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTC

```
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
|||||||  |||||
GTGGAGGATGAGGAT 94
```

050/c
, Application US/09849626
o. US20020197669A1
MATION:
angur, Chaitanya
Fanger, Gary
Wang, Aijun
Wang, Tongtong
Switzer, Anne
McNeill, Patricia
Clapper, Jonathan
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF LUNG CANCER
CE: 210121.478C16
ICATION NUMBER: US/09/849,626
NG DATE: 2001-05-03
Q ID NOS: 1926
stSEQ for Windows Version 3.0
0

omo sapien
050

```
          42.6%;  Score 32.8;  DB 9;  Length 472;
nilarity   64.5%;  Pred. No. 0.018;
  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| ||  || |     |         |  |  | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
|||||||  |||||
GTGGAGGATGAGGAT 94
```

050/c
, Application US/10017754
o. US20030054363A1
MATION:
enderson, Robert A.
Wang, Tongtong
Watanabe, Yoshihiro
Johnson, Jeffrey C.
Retter, Marc W.
Marnerakis, Margarita
Carter, Darrick
Fanger, Gary R.
Vedvick, Thomas S.
Bangur, Chaitanya S.
McNabb, Andria
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
ENTION: AND DIAGNOSIS OF LUNG CANCER
CE: 210121.478C18
ICATION NUMBER: US/10/017,754
NG DATE: 2001-10-29
Q ID NOS: 2004
stSEQ for Windows Version 4.0
0

omo sapiens
050

```
  Query Match           42.6%;  Score 32.8;  DB 14;  Length 472;
  Best Local Similarity   64.5%;  Pred. No. 0.018;
  Matches   49;  Conservative    0;  Mismatches   27;  Indels    0;

Qy          2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              ||| |||||||||| ||| ||  || |     |         |  |  | ||||
Db        169 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy         62 CGTGGAGGACAAGGAT 77
              |||||||||  |||||
Db        109 CGTGGAGGATGAGGAT 94
```

RESULT 39
US-10-113-872-1050/c
; Sequence 1050, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1050
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-1050

```
  Query Match           42.6%;  Score 32.8;  DB 14;  Length 472;
  Best Local Similarity   64.5%;  Pred. No. 0.018;
  Matches   49;  Conservative    0;  Mismatches   27;  Indels    0;

Qy          2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              ||| |||||||||| ||| ||  || |     |         |  |  | ||||
Db        169 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy         62 CGTGGAGGACAAGGAT 77
              |||||||||  |||||
Db        109 CGTGGAGGATGAGGAT 94
```

RESULT 40
US-09-920-300A-435/c
; Sequence 435, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 435
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-435

42.6%; Score 32.8; DB 9; Length 476;
milarity 64.5%; Pred. No. 0.018;
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| ||  || |     |        |  |  | ||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
|||||||  |||||
GTGGAGGATGAGGAT 94
```

d: April 13, 2004, 21:46:10
5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 18:19:17 ; Search time 43 Seconds
(without alignments)
993.749 Million cell updates/sec

US-10-090-326-24
77
1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

682709 seqs, 277475446 residues

hits satisfying chosen parameters: 1365418

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

is the number of results predicted by chance to have a
ater than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

	% Query Match	Length	DB	ID	Description
	58.4	862	3	US-08-468-856B-4	Sequence 4, Appli
	58.4	862	3	US-08-468-859A-4	Sequence 4, Appli
	58.4	2839	3	US-08-468-856B-5	Sequence 5, Appli
	58.4	2839	3	US-08-468-859A-5	Sequence 5, Appli
	56.4	2031	1	US-08-217-299-2	Sequence 2, Appli
	56.4	2097	2	US-08-602-725-35	Sequence 35, Appl
	56.4	2220	1	US-08-389-459A-16	Sequence 16, Appl
	56.4	2220	3	US-08-987-867A-16	Sequence 16, Appl
	56.4	2349	2	US-08-184-009-145	Sequence 145, App
	56.4	2349	2	US-08-458-356-145	Sequence 145, App
	56.4	2349	3	US-08-460-736-145	Sequence 145, App
	56.4	2349	4	US-09-535-370-145	Sequence 145, App
	56.4	2434	2	US-08-184-009-144	Sequence 144, App
	56.4	2434	2	US-08-458-356-144	Sequence 144, App
	56.4	2434	3	US-08-460-736-144	Sequence 144, App
	56.4	2434	4	US-09-535-370-144	Sequence 144, App
	53.2	588	3	US-09-385-982-447	Sequence 447, App
	42.6	472	4	US-09-702-705-1050	Sequence 1050, Ap
	42.6	472	4	US-09-736-457-1050	Sequence 1050, Ap
	42.6	472	4	US-09-614-124B-1050	Sequence 1050, Ap
	42.6	472	4	US-09-671-325-1050	Sequence 1050, Ap
	42.6	534	4	US-09-702-705-1325	Sequence 1325, Ap
	42.6	534	4	US-09-736-457-1325	Sequence 1325, Ap
	42.6	534	4	US-09-614-124B-1325	Sequence 1325, Ap
	42.6	534	4	US-09-671-325-1325	Sequence 1325, Ap
	42.6	536	4	US-09-702-705-1461	Sequence 1461, Ap
	42.6	536	4	US-09-736-457-1461	Sequence 1461, Ap
c	28	32.8	42.6	536	4	US-09-614-124B-1461	Sequenc
c	29	32.8	42.6	536	4	US-09-671-325-1461	Sequenc
	30	32.8	42.6	547	3	US-09-385-982-424	Sequenc
	31	32.8	42.6	613	3	US-09-385-982-194	Sequenc
	32	32.8	42.6	652	3	US-09-385-982-337	Sequenc
	33	32.8	42.6	656	3	US-09-385-982-316	Sequenc
	34	28	36.4	3096	4	US-09-489-039A-4015	Sequenc
	35	28	36.4	3152	1	US-07-924-028A-3	Sequenc
	36	28	36.4	3365	1	US-07-789-915A-1	Sequenc
	37	28	36.4	3365	1	US-08-005-002C-1	Sequenc
	38	28	36.4	3365	1	US-08-487-203A-1	Sequenc
	39	28	36.4	4279	4	US-09-446-402A-1	Sequenc
	40	28	36.4	4810	3	US-08-852-629-11	Sequenc
	41	28	36.4	4838	3	US-08-852-629-15	Sequenc
	42	28	36.4	6700	4	US-09-654-449-1	Sequenc
	43	28	36.4	7252	4	US-09-238-356-27	Sequenc
	44	28	36.4	7366	6	5169760-3	Patent N
c	45	28	36.4	7387	4	US-09-238-356-28	Sequenc

ALIGNMENTS

RESULT 1
US-08-468-856B-4
; Sequence 4, Application US/08468856B
; Patent No. 6013772
; GENERAL INFORMATION:
; APPLICANT: Barnett, Thomas; Elting, James; Kamarck, Michael;
; APPLICANT: Kretschmer, Axel
; TITLE OF INVENTION: CDNAS CODING FOR MEMBERS OF THE
; TITLE OF INVENTION: CARCINOEMBRYONIC ANTIGEN FAMILY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Horn Kramer & Woods
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 Mb storage
; COMPUTER: APPLE MACINTOSH 6500
; OPERATING SYSTEM: SYSTEM 7.5
; SOFTWARE: WordPerfect 3.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,856B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,974
; FILING DATE: 08-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/760,031
; FILING DATE: 13-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/274,107
; FILING DATE: 21-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/207,678
; FILING DATE: 16-JUN-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/060,031
; FILING DATE: 19-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/016,683
; FILING DATE: 19-FEB-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/896,361
; FILING DATE: 13-AUG-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141

```
'E/DOCKET NUMBER:  MDI 242.10-KGB
IICATION INFORMATION:
IE:  (914) 332-1700
   (914) 332-1844
 FOR SEQ ID NO:  4:
:HARACTERISTICS:
  862 nucleotides
ucleic acid
)NESS:  single
/:  linear
·4

           58.4%;  Score 45;  DB 3;  Length 862;
.milarity   74.0%;  Pred. No. 8.7e-08;
   Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||| |   |      | |    || |  | ||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 591

'CGTGGAGGACAAGGAT 77
|||||||||||||||
'CGTGGAGGACAAGGAT 608


4
pplication US/08468859A
22958
RMATION:
  Barnett, Thomas; Elting, James; Kamarck, Michael;
  Kretschmer, Axel
NVENTION:  CDNAS CODING FOR MEMBERS OF THE
NVENTION:  CARCINOEMBRYONIC ANTIGEN FAMILY
SEQUENCES:  14
ENCE ADDRESS:
E:  Sprung Horn Kramer & Woods
 660 White Plains Road
'arrytown
New York
  USA
591-5144
EADABLE FORM:
'YPE:  Diskette, 3.50 inch, 2.0 Mb storage
:  APPLE MACINTOSH 6500
G SYSTEM:  SYSTEM 7.5
:  WordPerfect 3.5
'PLICATION DATA:
ION NUMBER:  US/08/468,859A
ATE:  06-JUN-1995
CATION:  435
ICATION DATA:
ION NUMBER:  US 08/027,974
ATE:  08-MAR-1993
ICATION DATA:
ION NUMBER:  US 07/760,031
ATE:  13-SEP-1991
ICATION DATA:
ION NUMBER:  US 07/274,107
ATE:  21-NOV-1988
ICATION DATA:
'ION NUMBER:  US 07/207,678
ATE:  16-JUN-1988
ICATION DATA:
'ION NUMBER:  US 07/060,031
ATE:  19-JUN-1987
ICATION DATA:
'ION NUMBER:  US 07/016,683
ATE:  19-FEB-1987
ICATION DATA:
'ION NUMBER:  US 06/896,361
ATE:  13-AUG-1986
GENT INFORMATION:
```

```
;       NAME:  Kurt G. Briscoe
;       REGISTRATION NUMBER:  33,141
;       REFERENCE/DOCKET NUMBER:  MDI 242.9-KGB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (914) 332-1700
;       TELEFAX:  (914) 332-1844
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  862 nucleotides
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-468-859A-4

  Query Match             58.4%;  Score 45;  DB 3;  Length 862;
  Best Local Similarity   74.0%;  Pred. No. 8.7e-08;
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||| |   |      | |    || |  | |||
Db        532 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        592 CCGTGGAGGACAAGGAT 608


RESULT 3
US-08-468-856B-5
; Sequence 5, Application US/08468856B
; Patent No. 6013772
;  GENERAL INFORMATION:
;    APPLICANT:  Barnett, Thomas; Elting, James; Kamarck, Michael;
;    APPLICANT:  Kretschmer, Axel
;    TITLE OF INVENTION:  CDNAS CODING FOR MEMBERS OF THE
;    TITLE OF INVENTION:  CARCINOEMBRYONIC ANTIGEN FAMILY
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sprung Horn Kramer & Woods
;      STREET:  660 White Plains Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-5144
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 2.0 Mb storage
;      COMPUTER:  APPLE MACINTOSH 6500
;      OPERATING SYSTEM:  SYSTEM 7.5
;      SOFTWARE:  WordPerfect 3.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,856B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/027,974
;      FILING DATE:  08-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/760,031
;      FILING DATE:  13-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/274,107
;      FILING DATE:  21-NOV-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/207,678
;      FILING DATE:  16-JUN-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/060,031
;      FILING DATE:  19-JUN-1987
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/016,683
;      FILING DATE:  19-FEB-1987
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/896,361
```

```
ATE:  13-AUG-1986
GENT INFORMATION:
urt G. Briscoe
TION NUMBER:  33,141
E/DOCKET NUMBER:  MDI 242.10-KGB
ICATION INFORMATION:
E:  (914) 332-1700
  (914) 332-1844
FOR SEQ ID NO:  5:
HARACTERISTICS:
 2839 nucleotides
ucleic acid
NESS:  single
:  linear
5

          58.4%;  Score 45;  DB 3;  Length 2839;
milarity   74.0%;  Pred. No. 1.3e-07;
  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| |   |     | |    || |  | ||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1538

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1555



5
pplication US/08468859A
22958
RMATION:
  Barnett, Thomas; Elting, James; Kamarck, Michael;
  Kretschmer, Axel
NVENTION:  CDNAS CODING FOR MEMBERS OF THE
NVENTION:  CARCINOEMBRYONIC ANTIGEN FAMILY
SEQUENCES:  14
ENCE ADDRESS:
E:  Sprung Horn Kramer & Woods
 660 White Plains Road
arrytown
New York
  USA
591-5144
EADABLE FORM:
YPE:  Diskette, 3.50 inch, 2.0 Mb storage
:  APPLE MACINTOSH 6500
G SYSTEM:  SYSTEM 7.5
:  WordPerfect 3.5
PLICATION DATA:
ION NUMBER:  US/08/468,859A
ATE:  06-JUN-1995
CATION:  435
ICATION DATA:
ION NUMBER:  US 08/027,974
ATE:  08-MAR-1993
ICATION DATA:
ION NUMBER:  US 07/760,031
ATE:  13-SEP-1991
ICATION DATA:
ION NUMBER:  US 07/274,107
ATE:  21-NOV-1988
ICATION DATA:
ION NUMBER:  US 07/207,678
ATE:  16-JUN-1988
ICATION DATA:
ION NUMBER:  US 07/060,031
ATE:  19-JUN-1987
ICATION DATA:
ION NUMBER:  US 07/016,683
ATE:  19-FEB-1987
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/896,361
;      FILING DATE:  13-AUG-1986
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurt G. Briscoe
;      REGISTRATION NUMBER:  33,141
;      REFERENCE/DOCKET NUMBER:  MDI 242.9-KGB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (914) 332-1700
;      TELEFAX:  (914) 332-1844
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2839 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-468-859A-5

  Query Match             58.4%;  Score 45;  DB 3;  Length 2839;
  Best Local Similarity   74.0%;  Pred. No. 1.3e-07;
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||| |   |     | |    || |  | |||
Db       1479 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||
Db       1539 CCGTGGAGGACAAGGAT 1555


RESULT 5
US-08-217-299-2
; Sequence 2, Application US/08217299
; Patent No. 5672513
;  GENERAL INFORMATION:
;    APPLICANT:  Mach, J. P.
;    APPLICANT:  Pelegrin, A.
;    APPLICANT:  Terskikh, A.
;    TITLE OF INVENTION:  Carcinoembryonic Antigen Derivatives
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/217,299
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 93810214.2
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pokras, Bruce A.
;      REGISTRATION NUMBER:  32,748
;      REFERENCE/DOCKET NUMBER:  RAN 4093/096
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-5801
;      TELEFAX:  (201) 235-3500
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
```

```
'YPE:  DNA (genomic)
'AL:  NO
':  NO
'OURCE:
1:  Homo sapiens
'

          56.4%;  Score 43.4;  DB 1;  Length 2031;
.milarity   72.7%;  Pred. No. 4.9e-07;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

'GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |     | |   ||  | | ||||||||||
'GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1534

'CGTGGAGGACAAGGAT 77
||||||||||||||||
'CGTGGAGGACAAGGAT 1551



5
Application US/08602725
65710
'RMATION:
  BODMER, WALTER F
  DURBIN, HELGA
  SNARY, DAVID
  STEWART, LORNA MD
  YOUNG, SUSAN
  BATES, PAUL A
NVENTION:  MONOCLONAL ANTIBODIES FOR USE IN
NVENTION:  DIAGNOSIS AND TREATMENT OF COLORECTAL CANCER
SEQUENCES:  40
ENCE ADDRESS:
E:  NIXON & VANDERHYE P.C.
 1100 NORTH GLEBE ROAD, 8TH FLOOR
RLINGTON
VA
  USA
201
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.30
PLICATION DATA:
ION NUMBER:  US/08/602,725
ATE:  02-FEB-1996
CATION:  530
ICATION DATA:
ION NUMBER:  PCT/GB94/01816
ATE:  19-AUG-1994
ICATION DATA:
ION NUMBER:  GB 9317423
ATE:  21-AUG-1993
GENT INFORMATION:
ADOFF, B.J.
TION NUMBER:  36663
E/DOCKET NUMBER:  1090-8
ICATION INFORMATION:
E:  703-816-4091
  703-816-4100
FOR SEQ ID NO:  35:
HARACTERISTICS:
 2097 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  cDNA
AL:  NO
OURCE:
:  Human carcinoembryonic antigen
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2094
US-08-602-725-35

  Query Match           56.4%;  Score 43.4;  DB 2;  Length 2097;
  Best Local Similarity  72.7%;  Pred. No. 4.9e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||||     |     | |   ||  | | |||
Db       1463 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              ||||||||||||||||
Db       1523 CCGTGGAGGACAAGGAT 1539


RESULT 7
US-08-389-459A-16
; Sequence 16, Application US/08389459A
; Patent No. 5817512
;  GENERAL INFORMATION:
;    APPLICANT:  Morrow, Casey D. and Porter, Donna, C.
;    TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT POLIOVIRUS
;    TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, SUITE 510
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/389,459A
;      FILING DATE:  15-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Silveri, Jean M.
;      REGISTRATICN NUMBER:  39,030
;      REFERENCE/DOCKET NUMBER:  UAG-004CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2220 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2203
US-08-389-459A-16

  Query Match           56.4%;  Score 43.4;  DB 1;  Length 2220;
  Best Local Similarity  72.7%;  Pred. No. 5e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||||     |     | |   ||  | | |||
```

```
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1630

CGTGGAGGACAAGGAT 77
||||||||||||||||
CGTGGAGGACAAGGAT 1647



16
Application US/08987867A
63384
RMATION:
  C. Morrow et al.
NVENTION:  ENCAPSIDATED RECOMBINANT VIRAL
NVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
NVENTION:  USING SAME
SEQUENCES:  23
ENCE ADDRESS:
E:  LAHIVE & COCKFIELD
 28 STATE STREET
OSTON
MASSACHUSETTS
  USA
109
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  ASCII
PLICATION DATA:
ION NUMBER:  US/08/987,867A
ATE:  09-DEC-1997
CATION:  424
ICATION DATA:
ION NUMBER:  US 08/087,009
ATE:  01-JUL-1993
ENT INFORMATION:
ers, Louis
TION NUMBER:  35,965
E/DOCKET NUMBER:  UAP-004CPDV
ICATION INFORMATION:
E:  (617) 227-7400
  (617) 742-4214
OR SEQ ID NO:  16:
HARACTERISTICS:
 2220 base pairs
ucleic acid
NESS:  single
  linear
YPE:  cDNA

  CDS
  1..2203
.6

          56.4%;  Score 43.4;  DB 3;  Length 2220;
ilarity   72.7%;  Pred. No. 5e-07;
 Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |   |  |    ||  |  | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1630

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1647



5
Application US/08184009
3975
MATION:
```

```
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/184,009
;      FILING DATE:  19-JAN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  145:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2349 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-184-009-145

  Query Match           56.4%;  Score 43.4;  DB 2;  Length 2349;
  Best Local Similarity  72.7%;  Pred. No. 5.1e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||     |   |  |    ||  |  | |||
Db       1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1718 CCGTGGAGGACAAGGAT 1734


RESULT 10
US-08-458-356-145
; Sequence 145, Application US/08458356
; Patent No. 5942235
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
PLICATION DATA:
ION NUMBER:  US/08/458,356
ATE:  02-JUN-1995
CATION:  424
ICATION DATA:
ION NUMBER:  US 08/184,009
ATE:  19-JAN-1994
GENT INFORMATION:
rommer, William S.
TION NUMBER:  25,506
E/DOCKET NUMBER:  454310-2530
ICATION INFORMATION:
E:  (212) 840-3333
  (212) 840-0712
425066CURTMS
FOR SEQ ID NO:  145:
HARACTERISTICS:
 2349 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  cDNA
45

          56.4%;  Score 43.4;  DB 2;  Length 2349;
milarity   72.7%;  Pred. No. 5.1e-07;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |  | |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1717

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1734



45
 Application US/08460736
65189
RMATION:
  Paoletti, Enzo
  Tartaglia, James
  Cox, William I.
NVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
SEQUENCES:  217
ENCE ADDRESS:
E:  Curtis, Morris & Safford
 530 Fifth Avenue
ew York
NY
  USA
036
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/460,736
ATE:  02-JUN-1995
CATION:  514
ICATION DATA:
ION NUMBER:  US 08/184,009
ATE:  19-JAN-1994
GENT INFORMATION:
rommer, William S.
TION NUMBER:  25,506
E/DOCKET NUMBER:  454310-2530
ICATION INFORMATION:
E:  (212) 840-3333
  (212) 840-0712
;       TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  145:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  2349 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-460-736-145

  Query Match            56.4%;  Score 43.4;  DB 3;  Length 2349;
  Best Local Similarity  72.7%;  Pred. No. 5.1e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||     |  | |    ||  |  | |||
Db       1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1718 CCGTGGAGGACAAGGAT 1734


RESULT 12
US-09-535-370-145
; Sequence 145, Application US/09535370
; Patent No. 6537594
;   GENERAL INFORMATION:
;        APPLICANT: Paoletti, Enzo
;                   Tartaglia, James
;                   Cox, William I.
;        TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
;        NUMBER OF SEQUENCES: 217
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Curtis, Morris & Safford
;             STREET: 530 Fifth Avenue
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/535,370
;             FILING DATE: 24-Mar-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/460,736
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Frommer, William S.
;             REGISTRATION NUMBER: 25,506
;             REFERENCE/DOCKET NUMBER: 454310-2530
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 840-3333
;             TELEFAX: (212) 840-0712
;             TELEX: 425066CURTMS
;   INFORMATION FOR SEQ ID NO: 145:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2349 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 145:
US-09-535-370-145

  Query Match            56.4%;  Score 43.4;  DB 4;  Length 2349;
  Best Local Similarity  72.7%;  Pred. No. 5.1e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;
```

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1717

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1734



14
 Application US/08184009
33975
RMATION:
  Paoletti, Enzo
  Tartaglia, James
  Cox, William I.
NVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
SEQUENCES:  217
NCE ADDRESS:
S:  Curtis, Morris & Safford
 530 Fifth Avenue
ew York
NY
  USA
36
ADABLE FORM:
PE:  Floppy disk
   IBM PC compatible
 SYSTEM:  PC-DOS/MS-DOS
   PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ON NUMBER:  US/08/184,009
TE:  19-JAN-1994
ATION:  435
ENT INFORMATION:
ommer, William S.
ION NUMBER:  25,506
/DOCKET NUMBER:  454310-2530
CATION INFORMATION:
:  (212) 840-3333
  (212) 840-0712
25066CURTMS
OR SEQ ID NO:  144:
ARACTERISTICS:
2434 base pairs
cleic acid
ESS:  single
  linear
PE:  cDNA
4

          56.4%;  Score 43.4;  DB 2;  Length 2434;
ilarity   72.7%;  Pred. No. 5.2e-07;
 Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    |  |    ||  |  | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1714

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1731



4
Application US/08458356
2235
MATION:
 Paoletti, Enzo
 Tartaglia, James
```

```
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,356
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/184,009
;      FILING DATE:  19-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  144:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2434 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-458-356-144

  Query Match             56.4%;  Score 43.4;  DB 2;  Length 2434;
  Best Local Similarity   72.7%;  Pred. No. 5.2e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              ||||||||||||||||||||     |    |  |    ||  |  | ||||
Db       1655 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1715 CCGTGGAGGACAAGGAT 1731


RESULT 15
US-08-460-736-144
; Sequence 144, Application US/08460736
; Patent No. 6265189
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/460,736
ATE:  02-JUN-1995
CATION:  514
ICATION DATA:
ION NUMBER:  US 08/184,009
ATE:  19-JAN-1994
GENT INFORMATION:
rommer, William S.
TION NUMBER:  25,506
E/DOCKET NUMBER:  454310-2530
ICATION INFORMATION:
E:  (212) 840-3333
  (212) 840-0712
425066CURTMS
FOR SEQ ID NO:  144:
HARACTERISTICS:
 2434 base pairs
ucleic acid
NESS:  single
:  linear
YPE:  cDNA
44

          56.4%;  Score 43.4;  DB 3;  Length 2434;
milarity   72.7%;  Pred. No. 5.2e-07;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |     | |    ||  | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1714

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1731



44
 Application US/09535370
37594
ORMATION:
ANT: Paoletti, Enzo
     Tartaglia, James
     Cox, William I.
OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
 OF SEQUENCES: 217
PONDENCE ADDRESS:
DDRESSEE: Curtis, Morris & Safford
TREET: 530 Fifth Avenue
ITY: New York
TATE: NY
OUNTRY: USA
IP: 10036
ER READABLE FORM:
EDIUM TYPE: Floppy disk
OMPUTER: IBM PC compatible
PERATING SYSTEM: PC-DOS/MS-DOS
OFTWARE: PatentIn Release #1.0, Version #1.25
T APPLICATION DATA:
PPLICATION NUMBER: US/09/535,370
ILING DATE: 24-Mar-2000
LASSIFICATION: <Unknown>
APPLICATION DATA:
PPLICATION NUMBER: 08/460,736
ILING DATE: <Unknown>
EY/AGENT INFORMATION:
AME: Frommer, William S.
EGISTRATION NUMBER: 25,506
EFERENCE/DOCKET NUMBER: 454310-2530
MMUNICATION INFORMATION:
ELEPHONE: (212) 840-3333
;             TELEFAX: (212) 840-0712
;             TELEX: 425066CURTMS
;   INFORMATION FOR SEQ ID NO: 144:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2434 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 144:
US-09-535-370-144

  Query Match             56.4%;  Score 43.4;  DB 4;  Length 2434;
  Best Local Similarity   72.7%;  Pred. No. 5.2e-07;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||     |     | |    ||  | | |||
Db       1655 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1715 CCGTGGAGGACAAGGAT 1731


RESULT 17
US-09-385-982-447
; Sequence 447, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 447
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(588)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-447

  Query Match             53.2%;  Score 41;  DB 3;  Length 588;
  Best Local Similarity   79.2%;  Pred. No. 2.6e-06;
  Matches   61;  Conservative    0;  Mismatches   15;  Indels    1;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              |||||||||||||||||||||     ||  | || |||  ||  |  | |||
Db        314 AGACAATCACAGTCTCTGCGGAGCTGCCAAGCCCTCCAT-CTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        373 CCGTGGAGGACAAGGAT 389


RESULT 18
US-09-702-705-1050/c
; Sequence 1050, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
```

Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick
Retter, Marc
Mannion, Jane
Fan, Liqun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF LUNG CANCER
CE: 210121.478C14
ICATION NUMBER: US/09/702,705
NG DATE: 2000-10-30
) ID NOS: 1833
istSEQ for Windows Version 3.0
)

omo sapien
)50

42.6%; Score 32.8; DB 4; Length 472;
nilarity 64.5%; Pred. No. 0.0033;
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| || || | | | | | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
||||||| |||||
GTGGAGGATGAGGAT 94

)50/c
Application US/09736457
)9448
MATION:
Wang, Tongtong
Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick
Retter, Marc
Mannion, Jane
Fan, Liqun
Wang, Aijun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF LUNG CANCER
CE: 210121.478C15
ICATION NUMBER: US/09/736,457
NG DATE: 2000-12-13
) ID NOS: 1864
istSEQ for Windows Version 3.0
)

omo sapien
)50

42.6%; Score 32.8; DB 4; Length 472;
nilarity 64.5%; Pred. No. 0.0033;
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| || || | | | | | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
||||||| |||||
GTGGAGGATGAGGAT 94

RESULT 20
US-09-614-124B-1050/c
; Sequence 1050, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1050
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-1050

Query Match 42.6%; Score 32.8; DB 4; Length 472;
Best Local Similarity 64.5%; Pred. No. 0.0033;
Matches 49; Conservative 0; Mismatches 27; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
||| |||||||||| ||| || || | | | | ||||
Db 169 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy 62 CGTGGAGGACAAGGAT 77
||||||||| |||||
Db 109 CGTGGAGGATGAGGAT 94

RESULT 21
US-09-671-325-1050/c
; Sequence 1050, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AN
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1050
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-1050

Query Match 42.6%; Score 32.8; DB 4; Length 472;
Best Local Similarity 64.5%; Pred. No. 0.0033;
Matches 49; Conservative 0; Mismatches 27; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT

```
|| |||||||||| ||| ||  || |      |         |  |  | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 110

GTGGAGGACAAGGAT 77
|||||||| |||||
GTGGAGGATGAGGAT 94
```

```
325
, Application US/09702705
04010
MATION:
Wang, Tongtong
Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick
Retter, Marc
Mannion, Jane
Fan, Liqun
ENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION:  DIAGNOSIS OF LUNG CANCER
E:  210121.478C14
ICATION NUMBER: US/09/702,705
NG DATE:  2000-10-30
) ID NOS:  1833
astSEQ for Windows Version 3.0
;
```

```
omo sapien
325

          42.6%;  Score 32.8;  DB 4;  Length 534;
ilarity   64.5%;  Pred. No. 0.0034;
  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
| |||||||||| ||| ||  || |      |         |  |  | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 425

GTGGAGGACAAGGAT 77
|||||||  |||||
GTGGAGGATGAGGAT 441
```

```
325
  Application US/09736457
09448
MATION:
Wang, Tongtong
Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick
Retter, Marc
Mannion, Jane
Fan, Liqun
Wang, Aijun
ENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION:  DIAGNOSIS OF LUNG CANCER
E:  210121.478C15
ICATION NUMBER: US/09/736,457
NG DATE:  2000-12-13
) ID NOS:  1864
astSEQ for Windows Version 3.0
;
```

```
;   ORGANISM: Homo sapien
US-09-736-457-1325

  Query Match           42.6%;  Score 32.8;  DB 4;  Length 534;
  Best Local Similarity   64.5%;  Pred. No. 0.0034;
  Matches   49;  Conservative    0;  Mismatches   27;  Indels    0;

Qy          2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              ||| |||||||||| ||| ||  || |      |         |  |  | ||||
Db        366 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy         62 CGTGGAGGACAAGGAT 77
              |||||||||  |||||
Db        426 CGTGGAGGATGAGGAT 441
```

```
RESULT 24
US-09-614-124B-1325
; Sequence 1325, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C9
;  CURRENT APPLICATION NUMBER: US/09/614,124B
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 1668
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1325
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-614-124B-1325

  Query Match           42.6%;  Score 32.8;  DB 4;  Length 534;
  Best Local Similarity   64.5%;  Pred. No. 0.0034;
  Matches   49;  Conservative    0;  Mismatches   27;  Indels    0;

Qy          2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT(
              ||| |||||||||| ||| ||  || |      |         |  |  | ||||
Db        366 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT(

Qy         62 CGTGGAGGACAAGGAT 77
              |||||||||  |||||
Db        426 CGTGGAGGATGAGGAT 441
```

```
RESULT 25
US-09-671-325-1325
; Sequence 1325, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY ANI
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
```

ING DATE: 2000-09-26
EQ ID NOS: 1825
FastSEQ for Windows Version 3.0
25
4

Homo sapien
1325

42.6%; Score 32.8; DB 4; Length 534;
imilarity 64.5%; Pred. No. 0.0034;
; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
||| |||||||||| ||| || || | | | | |||||||| ||
GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 425

CGTGGAGGACAAGGAT 77
||||||||| |||||
CGTGGAGGATGAGGAT 441

:461/c
., Application US/09702705
:04010
:MATION:
Wang, Tongtong
Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick
Retter, Marc
Mannion, Jane
Fan, Liqun
ENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
ENTION: DIAGNOSIS OF LUNG CANCER
CE: 210121.478C14
ICATION NUMBER: US/09/702,705
NG DATE: 2000-10-30
Q ID NOS: 1833
astSEQ for Windows Version 3.0
1

omo sapien
461

42.6%; Score 32.8; DB 4; Length 536;
nilarity 64.5%; Pred. No. 0.0034;
Conservative 0; Mismatches 27; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| || || | | | | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 111

GTGGAGGACAAGGAT 77
||||||| |||||
GTGGAGGATGAGGAT 95

461/c
Application US/09736457
09448
MATION:
Wang, Tongtong
Bangur, Chaitanya S.
Lodes, Michael A.
Fanger, Gary
Vedvick, Tom
Carter, Darrick

; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY A
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1461
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-1461

Query Match 42.6%; Score 32.8; DB 4; Length 536;
Best Local Similarity 64.5%; Pred. No. 0.0034;
Matches 49; Conservative 0; Mismatches 27; Indels 0

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||| |||||||||| ||| || || | | | | |||
Db 170 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAAC

Qy 62 CGTGGAGGACAAGGAT 77
||||||||| |||||
Db 110 CGTGGAGGATGAGGAT 95

RESULT 28
US-09-614-124B-1461/c
; Sequence 1461, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1461
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-1461

Query Match 42.6%; Score 32.8; DB 4; Length 536;
Best Local Similarity 64.5%; Pred. No. 0.0034;
Matches 49; Conservative 0; Mismatches 27; Indels 0;

Qy 2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
||| |||||||||| ||| || || | | | | |||||
Db 170 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACT

Qy 62 CGTGGAGGACAAGGAT 77
||||||||| |||||
Db 110 CGTGGAGGATGAGGAT 95

RESULT 29
US-09-671-325-1461/c
; Sequence 1461, Application US/09671325
; Patent No. 6667154

```
)RMATION:
  Wang, Tongtong
  Bangur, Chaitanya S.
  Lodes, Michael A.
  Fanger, Gary
  Vedvick, Tom
  Carter, Darrick
  Retter, Marc
  Mannion, Jane
  Fan, Liqun
IVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
IVENTION:  DIAGNOSIS OF LUNG CANCER
:NCE:  210121.478C12
'LICATION NUMBER: US/09/671,325
:ING DATE:  2000-09-26
:EQ ID NOS:  1825
FastSEQ for Windows Version 3.0
61
6

Homo sapien
1461

            42.6%;  Score 32.8;  DB 4;  Length 536;
:imilarity   64.5%;  Pred. No. 0.0034;
;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
||| |||||||||| ||| ||  || |     |         | |  | ||||||| ||
GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 111

CGTGGAGGACAAGGAT 77
|||||||||  |||||
CGTGGAGGATGAGGAT 95



424
, Application US/09385982
262334
RMATION:
ENDEGE, WILSON O., ET AL.
VENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
VENTION: PRODUCTS:  II
NCE: CCDNA-260XX
LICATION NUMBER: US/09/385,982
ING DATE: 1999-08-30
LICATION NUMBER: 09/328,111
ING DATE: 1999-06-08
LICATION NUMBER: 60/117,393
ING DATE: 1999-01-27
LICATION NUMBER: 60/098,639
ING DATE: 1998-08-31
EQ ID NOS: 544
astSEQ for Windows Version 3.0
4
7

Homo sapiens

misc_feature
(1)...(547)
RMATION: n = A,T,C or G
424

            42.6%;  Score 32.8;  DB 3;  Length 547;
imilarity   64.5%;  Pred. No. 0.0034;
;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
||| |||||||||| ||| ||  || |     |         | |  | ||||||| ||
GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 196

Qy          62 CGTGGAGGACAAGGAT 77
               |||||||||  |||||
Db         197 CGTGGAGGATGAGGAT 212


RESULT 31
US-09-385-982-194
; Sequence 194, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 194
;   LENGTH: 613
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(613)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-194

  Query Match            42.6%;  Score 32.8;  DB 3;  Length 613;
  Best Local Similarity  64.5%;  Pred. No. 0.0036;
  Matches   49;  Conservative    0;  Mismatches   27;  Indels    0

Qy           2 GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
               ||| |||||||||| ||| ||  || |     |         | |  | |||
Db         140 GACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAAC

Qy          62 CGTGGAGGACAAGGAT 77
               |||||||||  |||||
Db         200 CGTGGAGGATGAGGAT 215


RESULT 32
US-09-385-982-337
; Sequence 337, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 337
;   LENGTH: 652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(652)
```

```
:MATION: n = A,T,C or G
:37

          42.6%;  Score 32.8;  DB 3;  Length 652;
.milarity   64.5%;  Pred. No. 0.0037;
  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

:ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| ||  || |     |          | |  | |||||||| ||
:ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 196

:GTGGAGGACAAGGAT 77
||||||||  |||||
:GTGGAGGATGAGGAT 212



16
 Application US/09385982
62334
MATION:
NDEGE, WILSON O., ET AL.
ENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
ENTION: PRODUCTS:  II
CE: CCDNA-260XX
ICATION NUMBER: US/09/385,982
NG DATE: 1999-08-30
ICATION NUMBER: 09/328,111
NG DATE: 1999-06-08
ICATION NUMBER: 60/117,393
NG DATE: 1999-01-27
ICATION NUMBER: 60/098,639
NG DATE: 1998-08-31
Q ID NOS: 544
stSEQ for Windows Version 3.0



omo sapiens

isc_feature
1)...(656)
MATION: n = A,T,C or G
16

          42.6%;  Score 32.8;  DB 3;  Length 656;
nilarity   64.5%;  Pred. No. 0.0037;
  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAACC 61
|| |||||||||| ||| ||  || |     |          | |  | |||||||| ||
ACGATCACAGTCTATGCAGAGCCACCCAAACCCTTCATCACCAGCAACAACTCCAACCC 196

GTGGAGGACAAGGAT 77
||||||||  |||||
GTGGAGGATGAGGAT 212



1015
  Application US/09489039A
:0836
MATION:
Gary Breton et. al
ENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
ENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
E:  2709.2004001
CATION NUMBER: US/09/489,039A
G DATE:  2000-01-27
ATION NUMBER:  US 60/117,747
 DATE:  1999-01-29
 ID NOS:  14342
;
```

```
;   LENGTH: 3096
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4015

  Query Match            36.4%;  Score 28;  DB 4;  Length 3096;
  Best Local Similarity  100.0%;  Pred. No. 0.44;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0

Qy         23 AGCATCATCCTCTGCATGGTCAGGTCAT 50
              ||||||||||||||||||||||||||||
Db       1097 AGCATCATCCTCTGCATGGTCAGGTCAT 1124


RESULT 35
US-07-924-028A-3
; Sequence 3, Application US/07924028A
; Patent No. 5470573
;  GENERAL INFORMATION:
;    APPLICANT:  Lubitz Werner, Szostak, Michael P.
;    TITLE OF INVENTION:  CARRIER-BOUND RECOMBINANT PROTEINS, PROCE
;    TITLE OF INVENTION:  FOR THE PRODUCTION AND USE AS IMMUNOGENS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/924,028A
;      FILING DATE:  30-SEP-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP91/00308
;      FILING DATE:  02-FEB-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 05 874
;      FILING DATE:  24-FEB-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5470573man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  HUBR 1027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3152 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-924-028A-3

  Query Match            36.4%;  Score 28;  DB 1;  Length 3152;
  Best Local Similarity  100.0%;  Pred. No. 0.45;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;

Qy         23 AGCATCATCCTCTGCATGGTCAGGTCAT 50
              ||||||||||||||||||||||||||||
Db       1088 AGCATCATCCTCTGCATGGTCAGGTCAT 1115


RESULT 36
US-07-789-915A-1
; Sequence 1, Application US/07789915A
; Patent No. 5212058
```

```
ORMATION:
:  Baker, Rohan T.
:  Tobias, John W.
:  Varshavsky, Alexander
INVENTION:  Ubiquitin-Specific Proteases
 SEQUENCES:  8
DENCE ADDRESS:
EE:  Hamilton, Brook, Smith & Reynolds, P.C.
  Two Militia Drive
Lexington
 Massachusetts
:  U.S.A.
2173
READABLE FORM:
TYPE:  Floppy disk
R:  IBM PC compatible
NG SYSTEM:  PC-DOS/MS-DOS
E:  PatentIn Release #1.0, Version #1.25
PPLICATION DATA:
TION NUMBER:  US/07/789,915A
DATE:  19911108
ICATION:  435
AGENT INFORMATION:
Brook, David E.
ATION NUMBER:  22,592
CE/DOCKET NUMBER:  MIT-5091AA
NICATION INFORMATION:
NE:  617-861-6240
:  617-861-9540
 FOR SEQ ID NO:  1:
CHARACTERISTICS:
  3365 base pairs
NUCLEIC ACID
DNESS:  double
:  linear

:  CDS
N:  1..3363
-1

          36.4%;  Score 28;  DB 1;  Length 3365;
milarity   100.0%;  Pred. No. 0.46;
  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

AGCATCATCCTCTGCATGGTCAGGTCAT 50
||||||||||||||||||||||||||||
AGCATCATCCTCTGCATGGTCAGGTCAT 1394


-1
pplication US/08005002C
94818
DRMATION:
  Baker, Rohan T.
  Tobias, John W.
  Varshavsky, Alexander
NVENTION:  Ubiquitin-Specific Proteases
SEQUENCES:  9
ENCE ADDRESS:
E:  Kevin M. Farrell, P.C.
  P.O. Box 999
ork Harbor
 Maine
  U.S.A.
911
EADABLE FORM:
YPE:  Floppy disk
:  IBM PC compatible
G SYSTEM:  PC-DOS/MS-DOS
:  PatentIn Release #1.0, Version #1.25
PLICATION DATA:
ION NUMBER:  US/08/005,002C
;      FILING DATE:  15-JAN-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/789,915
;      FILING DATE:  08-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Farrell, Kevin M.
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  MIT-5091AAZ
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  207-363-0558
;      TELEFAX:  207-363-0528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3363
US-08-005-002C-1

  Query Match            36.4%;  Score 28;  DB 1;  Length 3365;
  Best Local Similarity  100.0%;  Pred. No. 0.46;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0

Qy         23 AGCATCATCCTCTGCATGGTCAGGTCAT 50
              ||||||||||||||||||||||||||||
Db       1367 AGCATCATCCTCTGCATGGTCAGGTCAT 1394


RESULT 38
US-08-487-203A-1
; Sequence 1, Application US/08487203A
; Patent No. 5683904
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Rohan T.
;    APPLICANT:  Tobias, John W.
;    APPLICANT:  Varshavsky, Alexander
;    TITLE OF INVENTION:  Ubiquitin-Specific Proteases
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell, P.C.
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  Maine
;      COUNTRY:  U.S.A.
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,203A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/005,002
;      FILING DATE:  15-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Farrell, Kevin M.
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  MIT-5091A3Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  207-363-0558
;      TELEFAX:  207-363-0528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
```

linear

CDS
1..3363

36.4%; Score 28; DB 1; Length 3365;
ilarity 100.0%; Pred. No. 0.46;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

:CATCATCCTCTGCATGGTCAGGTCAT 50
||||||||||||||||||||||||||
:CATCATCCTCTGCATGGTCAGGTCAT 1394

plication US/09446402A
3003
ATION:
ack Jr., Charles A.
NTION: COMPOSITIONS AND METHODS FOR ACTIVATING
NTION: GENES OF INTEREST
E: 5722-2(35722/191928)
CATION NUMBER: US/09/446,402A
G DATE: 1999-12-20
TION NUMBER: PCT/US98/13093
DATE: 1998-06-24
TION NUMBER: 60/050,772
DATE: 1997-06-25
ID NOS: 19
tSEQ for Windows Version 4.0

tificial Sequence

ATION: Recombinant Molecule containing multiple cloning
ATION: site, kozak sequence, LacZ gene.
sc_feature
)...(64)
ATION: Multiple cloning site
sc_feature
5)...(79)
ATION: Consensus sequence for the "Kozak sequence"
ATION: (translation initiation)
im_transcript
0)...(4279)
ATION: Beta galactosidase

36.4%; Score 28; DB 4; Length 4279;
ilarity 100.0%; Pred. No. 0.5;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CATCATCCTCTGCATGGTCAGGTCAT 50
||||||||||||||||||||||||||
CATCATCCTCTGCATGGTCAGGTCAT 1864

pplication US/08852629
6825
MATION:
Moyer, Richard W
Li, Yi
Hall, Richard L
VENTION: ENTOMOPOXVIRUS-VERTEBRATE GENE DELIVERY
VENTION: VECTOR AND METHOD
EQUENCES: 17
NCE ADDRESS:
: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,629
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: UF-184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 407-426-7500
; TELEFAX: 407-839-8589
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-852-629-11

Query Match 36.4%; Score 28; DB 3; Length 4810;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 28; Conservative 0; Mismatches 0; Indels 0;

Qy 23 AGCATCATCCTCTGCATGGTCAGGTCAT 50
||||||||||||||||||||||||||
Db 2088 AGCATCATCCTCTGCATGGTCAGGTCAT 2115

Search completed: April 13, 2004, 20:16:04
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 17:09:09 ; Search time 217.5 Seconds
(without alignments)
1503.960 Million cell updates/sec

US-10-090-326-24
77
1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

3373863 seqs, 2124099041 residues

hits satisfying chosen parameters: 6747726

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

is the number of results predicted by chance to have a
ater than or equal to the score of the result being printed,
rived by analysis of the total score distribution.

SUMMARIES

Query Match	Length	DB	ID	Description
100.0	77	6	ABQ82558	Abq82558 CEA mimic
58.7	1943	4	AAH33528	Aah33528 Human col
58.4	862	1	AAN81583	Aan81583 LV7 cDNA
58.4	862	2	AAQ54351	Aaq54351 Carcinoem
58.4	862	2	AAV70152	Aav70152 Human CEA
58.4	1623	6	AAD39107	Aad39107 Human lun
58.4	2839	1	AAN81584	Aan81584 1LV7 cDNA
58.4	2839	1	AAN92780	Aan92780 cDNA sequ
58.4	2839	2	AAQ54352	Aaq54352 Carcinoem
58.4	2839	2	AAT46062	Aat46062 Carcinoem
57.9	2459	3	AAC77897	Aac77897 Human can
56.4	111	4	AAS57592	Aas57592 cDNA #268
56.4	190	4	AAS57578	Aas57578 cDNA #254
56.4	326	4	AAS58436	Aas58436 cDNA #111
56.4	326	4	AAS58714	Aas58714 cDNA #139
56.4	326	4	AAS58211	Aas58211 cDNA #887
56.4	532	6	ABV87469	Abv87469 Human col
56.4	571	4	AAS58049	Aas58049 cDNA #725
56.4	572	4	AAS58448	Aas58448 cDNA #112
56.4	2019	6	ABK86206	Abk86206 cDNA enco
56.4	2031	2	AAQ71567	Aaq71567 Carcinoem
56.4	2059	2	AAT36495	Aat36495 Immunogen
56.4	2097	2	AAQ82807	Aaq82807 Carcinoem
24	43.4	56.4	2105	6	AAI72497	Aai72497
25	43.4	56.4	2106	4	AAH20121	Aah20121
26	43.4	56.4	2106	5	AAD07347	Aad07347
27	43.4	56.4	2106	6	AAI72489	Aai72489
28	43.4	56.4	2106	9	ADE13860	Ade13860
29	43.4	56.4	2109	6	ABL54023	Abl54023
30	43.4	56.4	2220	2	AAT33302	Aat33302
31	43.4	56.4	2349	2	AAQ67869	Aaq67869
32	43.4	56.4	2349	2	AAZ08470	Aaz08470
33	43.4	56.4	2389	4	AAF61687	Aaf61687
34	43.4	56.4	2434	2	AAQ67868	Aaq67868
35	43.4	56.4	2434	2	AAZ08469	Aaz08469
36	43.4	56.4	2907	9	ADD78270	Add78270
37	43.4	56.4	2928	1	AAN81611	Aan81611
38	43.4	56.4	2928	2	AAT75431	Aat75431
39	43.4	56.4	2974	6	ABL64746	Abl64746
40	43.4	56.4	2974	6	ABN95819	Abn95819
41	43.4	56.4	2974	7	ABX76144	Abx76144
42	43.4	56.4	2974	7	ABX76396	Abx76396
43	43.4	56.4	2974	7	ADA84058	Ada84058
44	43.4	56.4	2974	7	ABQ83855	Abq83855
45	43.4	56.4	2974	8	ACF35963	Acf35963

ALIGNMENTS

RESULT 1
ABQ82558
ID ABQ82558 standard; DNA; 77 BP.
XX
AC ABQ82558;
XX
DT 18-DEC-2002 (first entry)
XX
DE CEA mimic internal control probe SEQ ID NO:24.
XX
KW Human; CEA; carcinoembryonic antigen; adenocarcinoma; oesophag
KW malignancy; probe; ss.
XX
OS Synthetic.
XX
PN WO200270751-A1.
XX
PD 12-SEP-2002.
XX
PF 04-MAR-2002; 2002WO-US006504.
XX
PR 02-MAR-2001; 2001US-0273277P.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Godfrey TE, Luketich JD, Raja S, Kelly LA, Finkelstein SD;
XX
DR WPI; 2002-732795/79.
XX
PT Multiplex PCR method for detecting malignancies, e.g. adenocar
PT the esophagus comprises conducting a PCR amplification on a DN
PT a PCR reaction mixture.
XX
PS Claim 47; Page 80; 141pp; English.
XX
CC The present invention describes a multiplex polymerase chain r
CC (PCR) (M1) comprising conducting PCR on a DNA sample in a react
CC mixture conducted in first and second amplification stages, ea
CC or more PCR cycles comprising denaturing, annealing and elongat
CC where the elongating step may be conducted at the same temperat
CC annealing step. The second amplification stage of (M1) is condu
CC different reaction conditions from that of the first amplificat
CC to modulate the relative rate of production of the first amplic
CC first primer set and a second amplicon by a second primer set
CC first and second amplification stages. Also described: (1) an
CC oligonucleotide comprising 15-28 bp or its derivative; (2) int

>stic; (3) rapid detection of a malignancy or of metastasised
inoma of the oesophagus; or (4) a cartridge for use in an
PCR system. (M1) is useful for detecting malignancies, e.g.
inoma of the oesophagus. (M1) eliminates contamination and
the time it takes to carry out a PCR reaction. The present
represents an internal control probe, which is used in an
rom the present invention

77 BP; 23 A; 21 C; 18 G; 15 T; 0 U; 0 Other;

100.0%; Score 77; DB 6; Length 77;
imilarity 100.0%; Pred. No. 4.3e-18;
: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 77

standard; cDNA; 1943 BP.

01 (first entry)

>n cancer antigen encoding cDNA SEQ ID NO:584.

.on cancer; colon cancer antigen; diagnosis; detection;
. carcinoma; ss.

>ns.

>0-A2.

>1.

>0; 2000WO-US026524.

>9; 99US-0157137P.
>9; 99US-0163280P.

JMAN GENOME SCI INC.

Barash SC, Birse CE, Rosen CA;

-235357/24.
AG74097.

>ids encoding 4277 human colon cancer-associated polypeptides,
> preventing, diagnosing and/or treating colorectal cancers.

>age 2637-2638; 9803pp; English.

>o AAH37195 and AAG73514 to AAG77788 represent human colon
sociated nucleic acid molecules (N) and proteins (P), where the
ire collectively known as colon cancer antigens. The colon
>igens have cytostatic activity and can be used in gene therapy
>e production. N and P may be used in the prevention, diagnosis
ment of diseases associated with inappropriate P expression. For
N and P may be used to treat disorders associated with decreased
> by rectifying mutations or deletions in a patient's genome
>t the activity of P by expressing inactive proteins or to
> the patients own production of P. Additionally, N may be used
> the colon cancer-associated Ps, by inserting the nucleic acids
st cell and culturing the cell to express the proteins. N and P
>d in the prevention, diagnosis and treatment of colorectal
s and cancers. AAH37196 to AAH37204 and AAB77789 represent

CC sequences used in the exemplification of the present invention
CC Pages 666 to 682 and page 7053 of the sequence listing were m
CC time of publication, meaning no sequences are present for SEQ
CC to 1052, 7921 and 7922
XX
SQ Sequence 1943 BP; 524 A; 597 C; 397 G; 422 T; 0 U; 3 Other;

Query Match 58.7%; Score 45.2; DB 4; Length 1943;
Best Local Similarity 72.7%; Pred. No. 3.5e-06;
Matches 56; Conservative 1; Mismatches 20; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| : | | | || | | ||
Db 1155 AGACAATCACAGTCTCTGCGGANSTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1215 CCGTGGAGGACAAGGAT 1231

RESULT 3
AAN81583
ID AAN81583 standard; cDNA; 862 BP.
XX
AC AAN81583;
XX
DT 25-MAR-2003 (revised)
DT 07-DEC-1990 (first entry)
XX
DE LV7 cDNA encoding carcinoembryonic antigen-A.
XX
KW Carcinoembryonic antigen-A; LV7; colorectal cancer; antibodie
KW immunoassay; probes; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3. .860
FT /*tag= a
FT /product= "carcinoembryonic antigen-A"
XX
PN EP263933-A.
XX
PD 20-APR-1988.
XX
PF 03-AUG-1987; 87EP-00111168.
XX
PR 13-AUG-1986; 86US-00896361.
PR 19-FEB-1987; 87US-00016683.
PR 19-JUN-1987; 87US-00060031.
XX
PA (MOLE-) MOLECULAR DIAGNOSTICS INC.
XX
PI Barnett TR, Elting JJ, Kamarck ME;
XX
DR WPI; 1988-106756/16.
DR P-PSDB; AAP81224.
XX
PT New nucleic acid sequences coding for carcinoembryonic antige
PT expressed polypeptide(s), useful for making diagnostic and the
PT antibodies.
XX
PS Claim 1 (a); Page 46; 68pp; English.
XX
CC The sequence can be expressed in lambda gt11. It has an interr
CC sequence of approx. 300 bases long and codes for a theoretical
CC 285 amino acids long. This cDNA encodes for sequences found ir
CC human genes and it is believed that this entire family of gene
CC for the CEA family of proteins. CEA is a marker for management
CC colorectal cancer. The antibodies are useful for immunoassay a
CC detection of CEA, and also when bonded to radionuclides or to>
CC CEA- expressing cells. The nucleic acid can be used as hybridi
CC probes. See also AAN81583-86. (Updated on 25-MAR-2003 to corre

dated on 25-MAR-2003 to correct PI field.)

2 BP; 232 A; 282 C; 180 G; 168 T; 0 U; 0 Other;

58.4%; Score 45; DB 1; Length 862;
ilarity 74.0%; Pred. No. 3.3e-06;
Conservative 0; Mismatches 20; Indels 0; Gaps 0;

:ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | | || | | ||||||||||
:ACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 591

:GTGGAGGACAAGGAT 77
|||||||||||||||
:GTGGAGGACAAGGAT 608

andard; cDNA; 862 BP.

(revised)
(first entry)

yonic antigen CEA-(a) coding sequence LV7.

yonic antigen; CEA; tumour marker; colorectal cancer;
; LoVo cell; ss.

s.

Location/Qualifiers
3. .860
/*tag= a
/note= "partial open reading frame, does not include
initiation or termination codons"

; 92US-00876821.

; 86US-00896361.
; 87US-00016683.
; 87US-00060031.
; 88US-00231741.

ECULAR DIAGNOSTICS INC.

Kamarck ME, Elting JJ;

06741/01.
54712.

d encoding carcinoembryonic antigens, - used in detection
diagnose different types of cancer.

g 1; 39pp; English.

a member of the carcinoembryonic antigen family. Sequence
an immunoreactive fragment of CEA contg. an internal repeat
prox. 300 bases long. The cDNA sequence was isolated from a
cell cDNA library as clone LV7 and can be used as a probe to
uences coding for other CEAs. (Updated on 25-MAR-2003 to
field.)

2 BP; 234 A; 280 C; 180 G; 168 T; 0 U; 0 Other;

58.4%; Score 45; DB 2; Length 862;
ilarity 74.0%; Pred. No. 3.3e-06;

Matches 57; Conservative 0; Mismatches 20; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
|||||||||||||||||||| | | | | || | | ||||
Db 532 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 592 CCGTGGAGGACAAGGAT 608

RESULT 5
AAV70152
ID AAV70152 standard; cDNA; 862 BP.
XX
AC AAV70152;
XX
DT 04-FEB-1999 (first entry)
XX
DE Human CEA protein immunoreactive fragment encoding cDNA.
XX
KW Human; CEA; carcinoembryonic antigen; immunoreactive; cloning
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3. .860
FT /*tag= a
FT /note= "no stop codon given"
XX
PN US5843761-A.
XX
PD 01-DEC-1998.
XX
PF 06-JUN-1995; 95US-00469987.
XX
PR 13-AUG-1986; 86US-00896361.
PR 19-FEB-1987; 87US-00016683.
PR 19-JUN-1987; 87US-00060031.
PR 12-AUG-1988; 88US-00231741.
PR 29-APR-1992; 92US-00876821.
PR 20-DEC-1993; 93US-00170134.
XX
PA (FARB) BAYER CORP FORMERLY MOLECULAR DIAGNOSTIC.
XX
PI Elting JJ, Kamarck ME, Barnett TR;
XX
DR WPI; 1999-044586/04.
DR P-PSDB; AAW83136.
XX
PT Cloning vectors - containing DNA inserts encoding carcinoembryc
PT antigen polypeptides.
XX
PS Example 8; Fig 1; 31pp; English.
XX
CC The present invention describes replicable cloning vectors cont
CC inserts encoding the carcinoembryonic antigen (CEA) polypeptide
CC by the DNA inserts of the plasmids contained in Escherichia col
CC ATCC 67169, ATCC 67312, ATCC 67710 and ATCC 67711. The cloning
CC are useful for cloning the DNA inserts. The present sequence re
CC DNA sequence which codes for an immunoreactive fragment of CEA
CC be expressed in lambda gt11
XX
SQ Sequence 862 BP; 233 A; 276 C; 185 G; 168 T; 0 U; 0 Other;

Query Match 58.4%; Score 45; DB 2; Length 862;
Best Local Similarity 74.0%; Pred. No. 3.3e-06;
Matches 57; Conservative 0; Mismatches 20; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
|||||||||||||||||||| | | | | || | | ||||
Db 532 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACT

```
CGTGGAGGACAAGGAT 77
||||||||||||||||
CGTGGAGGACAAGGAT 608
```

tandard; cDNA; 1623 BP.

2 (first entry)

-specific nucleic acid (LSNA) #14.

g-specific nucleic acid; LSNA; lung-specific protein; LSP;
r; non-cancerous disease; gene therapy; transgenic animal;
; engineered lung tissue; cytostatic; gene; ss.

ns.

```
      Location/Qualifiers
      3. .684
      /*tag= a
      /product= "Human LSP #14"
      /transl_except= (pos:33. .34, aa:His)
      /transl_except= (pos:335. .336, aa:Leu)
      /note= "These codons have an apparent deletion of 1
      nucleotide base each, which alters the reading frame; CDS
      does not include start codon"
      /partial
```

3-A2.

2.

1; 2001WO-US045180.

0; 2000US-0252055P.
0; 2000US-0252496P.

ADEXUS INC.

Recipon H, Chen S, Sun Y, Liu C, Turner LR;

471623/50.
E24334.

pecific nucleic acids, useful for identifying, diagnosing,
, staging, imaging, and treating lung cancer and non-cancerous
ates in lung tissues.

age 156-157; 187pp; English.

ion relates to new lung-specific nucleic acids (LSNA) and lung-
roteins (LSP). The lung-specific nucleic acids, polypeptides
itions comprising them are useful: for identifying, diagnosing,
, staging, imaging, and treating lung cancer and non-cancerous
ates in lung tissue; for identifying lung tissue; for
, identifying and/or designing agonists and antagonists of the
es; in gene therapy; in producing transgenic animals and cells;
ing engineered lung tissue for treatment and research; and as
n an array or computer program for pattern recognition of lung
The nucleic acids may be used as hybridisation probes to
aracterise and quantify hybridising nucleic acids in and
bridising nucleic acids from, both genomic and transcript-
cleic acid samples. The present sequence is human lung-specific

623 BP; 494 A; 419 C; 306 G; 403 T; 0 U; 1 Other;

58.4%; Score 45; DB 6; Length 1623;
milarity 74.0%; Pred. No. 3.9e-06;

```
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||| |   |     | |    ||  |  | |||
Db         51 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        111 CCGTGGAGGACAAGGAT 127
```

RESULT 7
AAN81584
ID AAN81584 standard; DNA; 2839 BP.
XX
AC AAN81584;
XX
DT 25-MAR-2003 (revised)
DT 07-DEC-1990 (first entry)
XX
DE 1LV7 cDNA encoding carcinoembryonic antigen-B.
XX
KW Carcinoembryonic antigen-B; 1LV7; colorectal cancer; antibodie
KW immunoassay; probes; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 5. .2110
FT /*tag= a
FT /product= "carcinoembryonic antigen-B"
XX
PN EP263933-A.
XX
PD 20-APR-1988.
XX
PF 03-AUG-1987; 87EP-00111168.
XX
PR 13-AUG-1986; 86US-00896361.
PR 19-FEB-1987; 87US-00016683.
PR 19-JUN-1987; 87US-00060031.
XX
PA (MOLE-) MOLECULAR DIAGNOSTICS INC.
XX
PI Barnett TR, Elting JJ, Kamarck ME;
XX
DR WPI; 1988-106756/16.
DR P-PSDB; AAP81222.
XX
PT New nucleic acid sequences coding for carcinoembryonic antigen
PT expressed polypeptide(s), useful for making diagnostic and the
PT antibodies.
XX
PS Claim 1 (b); Page 46; 68pp; English.
XX
CC CEA is a marker for management of colorectal cancer. The antib
CC useful for immunoassay and in vivo detection of CEA, and also
CC to radionuclides or toxins to kill CEA- expressing cells. The
CC acid can be used as hybridisation probes. See also AAN81583-86
CC on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 t
CC PI field.)
XX
SQ Sequence 2839 BP; 793 A; 815 C; 567 G; 664 T; 0 U; 0 Other;

```
  Query Match            58.4%;  Score 45;  DB 1;  Length 2839;
  Best Local Similarity  74.0%;  Pred. No. 4.5e-06;
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||| |   |     | |    ||  |  | |||
Db       1479 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
```

```
||||||||||||||||
CGTGGAGGACAAGGAT 1555
```

tandard; DNA; 2839 BP.

3 (revised)
2 (revised)
0 (first entry)

nce encoding CEA complete antigen 15LV7 (CEA-(b)).

ryonic antigens (CEA's); antibody (anti-CEA) production;

Location/Qualifiers
5. .2110

9.

9; 89EP-00110096.

3; 88US-00207678.
3; 88US-00274107.

LECULAR DIAGNOSTICS INC.

Elting JJ, Kamarck ME, Kretschmer A;

372000/51.
?93999.

for members of carcinoembryonic antigen family - used to
obes to detect antigen or in prodn. of polypeptide(s) and

Page 20-26; 65pp; English.

ice encodes the peptide sequence of the CEA-(b) antigen (pc
is used as a nucleic acid probe to detect the presence of CEA
a related nucleic acid in a test sample. Since different CEA's
in different cancers the presence of a particular CEA can act
ostic tool. (Updated on 31-OCT-2002 to add missing OS field.)
25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
field.)

39 BP; 790 A; 810 C; 575 G; 664 T; 0 U; 0 Other;

58.4%; Score 45; DB 1; Length 2839;
ilarity 74.0%; Pred. No. 4.5e-06;
Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||| |   |     | |   ||  |  | ||||||||||
ACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1538

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1555
```

andard; cDNA; 2839 BP.

AC AAQ54352;
XX
DT 25-MAR-2003 (revised)
DT 23-JUN-1994 (first entry)
XX
DE Carcinoembryonic antigen CEA-(b) coding sequence.
XX
KW Carcinoembryonic antigen; CEA; tumour marker; colorectal cance
KW immunoassay; LoVo cell; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 5. .2113
FT /*tag= a
FT /product= "CEA-(b)"
FT /transl_except= pos:182. .184,aa:Glu
FT /transl_except= pos:350. .352,aa:Asn
FT /transl_except= pos:1400. .1402,aa:Asn
FT mat_peptide 107. .2110
FT /*tag= b
XX
PN US5274087-A.
XX
PD 28-DEC-1993.
XX
PF 29-APR-1992; 92US-00876821.
XX
PR 13-AUG-1986; 86US-00896361.
PR 19-FEB-1987; 87US-00016683.
PR 19-JUN-1987; 87US-00060031.
PR 12-AUG-1988; 88US-00231741.
XX
PA (MOLE-) MOLECULAR DIAGNOSTICS INC.
XX
PI Barnett TR, Kamarck ME, Elting JJ;
XX
DR WPI; 1994-006741/01.
DR P-PSDB; AAR54713.
XX
PT Nucleic acid encoding carcinoembryonic antigens, - used in det
PT methods to diagnose different types of cancer.
XX
PS Claim 2; Fig 5; 39pp; English.
XX
CC CEA-(b) is a member of the carcinoembryonic antigen family. Th
CC sequence AAQ54352 was isolated from a LoVo tumour cell cDNA lil
CC Plasmid pc1LV7 comprising the cDNA insert sequence was deposit
CC E.coli as ATCC # 67312. (Updated on 25-MAR-2003 to correct PF
XX
SQ Sequence 2839 BP; 790 A; 811 C; 574 G; 664 T; 0 U; 0 Other;

Query Match 58.4%; Score 45; DB 2; Length 2839;
Best Local Similarity 74.0%; Pred. No. 4.5e-06;
Matches 57; Conservative 0; Mismatches 20; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              |||||||||||||||||||| |   |     | |   ||  |  | ||||
Db       1479 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1539 CCGTGGAGGACAAGGAT 1555
```

RESULT 10
AAT46062
ID AAT46062 standard; cDNA; 2839 BP.
XX
AC AAT46062;
XX
DT 25-MAR-2003 (revised)
DT 17-MAR-1997 (first entry)

ryonic antigen CEA-b cDNA clone c1LV7.

ryonic antigen; CEA-b; colorectal cancer; diagnosis; therapy;

ens.

Location/Qualifiers
5. .2113
/*tag= a
/transl_except= pos:182. .184:_aa:Glu
/transl_except= pos:350. .352:_aa:Asn
/transl_except= pos:1400. .1402:_aa:Asn

-A.

96.

93; 93US-00170134.

86; 86US-00896361.
87; 87US-00016683.
87; 87US-00060031.
88; 88US-00231741.
92; 92US-00876821.

AYER CORP.

, Elting JJ, Kamarck ME;

-505404/50.
AW06872.

ing carcinoembryonic antigen polypeptide - for diagnostic or
ic use.

; Fig 5A-G; 32pp; English.

one (AAT46062), designated c1LV7, encodes carcinoembryonic
EA-b (AAW06872). It was identified in LoVo cDNA using a
ed CEA clone LV7 (see also AAT46061) fragment probe. A plasmid
. contg. pc1LV7 has been deposited as ATCC 67312. CEA gene
(see also AAT46063-64) can be used as probes to isolate other
the CEA gene family, to determine expression of CEA genes in
mour types, and to produce recombinant CEA polypeptides, esp.
seful in immunoassays. (Updated on 25-MAR-2003 to correct PF

839 BP; 790 A; 811 C; 574 G; 664 T; 0 U; 0 Other;

58.4%; Score 45; DB 2; Length 2839;
milarity 74.0%; Pred. No. 4.5e-06;
Conservative 0; Mismatches 20; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1538

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 1555

standard; cDNA; 2459 BP.

01 (first entry)

cer associated gene sequence SEQ ID NO:291.

XX
KW Human; cancer associated gene; cancer antigen; detection; can
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodula
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; an
KW antiinflammatory; antithyroid; antiallergic; antibacterial; c
KW dermatological; neuroprotective; thrombolytic; coagulant; noo
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; infl
KW immune disorder; haematopoietic cell disorder; autoimmune dis
KW allergic reaction; graft versus host disease; organ rejection
KW haemostatic; thrombolytic; cardiovascular disorder; infection
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR P-PSDB; AAB43688.
XX
PT Novel isolated nucleic acids comprising sequences encoding pep
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 1; Page 852; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated prote
CC AAB43398 to AAB44239. The proteins can have activities based
CC tissues and cells the genes are expressed in. Example of acti
CC include: cytostatic; proliferative; vulnerary; immunomodulato
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; a
CC dermatological; neuroprotective; cardiant; thrombolytic; coag
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing,
CC ameliorating medical conditions and diagnosing pathological c
CC Polynucleotides, polypeptides, antibodies, agonists and antag
CC the present invention may be used to treat immune disorders b
CC or inhibiting the proliferation, differentiation or mobilisat
CC immune cells, to treat disorders of haematopoietic cells, aut
CC disorders, allergic reactions, graft versus host disease and
CC rejection, modulate haemostatic or thrombolytic activity, mod
CC inflammation, cancers, cardiovascular disorders, neurological
CC bacterial or viral infections. The peptides, nucleotides, ant
CC agonists and antagonists may be also be used in drug screens.
CC AAC78457 and AAB44240 represent sequences used in the exempli
CC the present invention
XX
SQ Sequence 2459 BP; 661 A; 748 C; 517 G; 518 T; 0 U; 15 Other;

Query Match 57.9%; Score 44.6; DB 3; Length 2459;
Best Local Similarity 72.7%; Pred. No. 6.1e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||| | | | || | | |||
Db 1582 AGACAATCACAGTCTCTGCGGANNTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1642 CCGTGGAGGACAAGGAT 1658

RESULT 12
AAS57592

tandard; cDNA; 111 BP.

2 (first entry)

encoding portion of a human colon tumour protein.

on tumour protein; colon cancer; gene therapy; cytostatic; ss.

ns.

7-A2.

1.

1; 2001WO-US009246.

0; 2000US-0191597P.
0; 2000US-0202024P.
0; 2000US-0202189P.

RIXA CORP.

, Xu J, King GE;

611627/70.

tumor proteins and related nucleic acid, useful for treatment,
, diagnosis and monitoring of cancer.

age 101; 299pp; English.

invention relates to the isolation of novel cDNA sequences
or at least an immunogenic portion of human colon tumour
The sequences of the invention are useful in pharmaceutical
ns and vaccines for the prevention and treatment of cancers
lon cancer. They are also useful for the diagnosis and
of such cancers. Antibodies to the colon tumour proteins and
esenting cells that express polynucleotides encoding colon
teins can be used to inhibit the development of cancers. T-
react specifically with colon tumour proteins are useful for
umour cells from samples (e.g. blood) and for cancer treatment.
cleotides sequences are also useful in gene therapy. AAS57325-
epresent the cDNA sequences of the invention that encode for
f human colon tumour proteins

11 BP; 31 A; 39 C; 24 G; 17 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 4; Length 111;
nilarity 72.7%; Pred. No. 7.2e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |   |  |   ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97
```

tandard; cDNA; 190 BP.

2 (first entry)

encoding portion of a human colon tumour protein.

on tumour protein; colon cancer; gene therapy; cytostatic; ss.

XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US009246.
XX
PR 24-MAR-2000; 2000US-0191597P.
PR 04-MAY-2000; 2000US-0202024P.
PR 05-MAY-2000; 2000US-0202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
DR WPI; 2001-611627/70.
XX
PT New colon tumor proteins and related nucleic acid, useful for
PT prevention, diagnosis and monitoring of cancer.
XX
PS Claim 4; Page 99; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA se
CC encoding for at least an immunogenic portion of human colon tu
CC proteins. The sequences of the invention are useful in pharmac
CC compositions and vaccines for the prevention and treatment of
CC such as colon cancer. They are also useful for the diagnosis a
CC monitoring of such cancers. Antibodies to the colon tumour pro
CC antigen presenting cells that express polynucleotides encoding
CC tumour proteins can be used to inhibit the development of canc
CC cells that react specifically with colon tumour proteins are u
CC removing tumour cells from samples (e.g. blood) and for cancer
CC The polynucleotides sequences are also useful in gene therapy.
CC AAS58880 represent the cDNA sequences of the invention that en
CC portions of human colon tumour proteins
XX
SQ Sequence 190 BP; 53 A; 60 C; 44 G; 33 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 190;
Best Local Similarity 72.7%; Pred. No. 8.3e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||     |   |  |   ||  |  | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97
```

RESULT 14
AAS58436
ID AAS58436 standard; cDNA; 326 BP.
XX
AC AAS58436;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #1112 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytos
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US009246.
XX

; 2000US-0191597P.
; 2000US-0202024P.
; 2000US-0202189P.

IXA CORP.

Xu J, King GE;

:11627/70.

umor proteins and related nucleic acid, useful for treatment, diagnosis and monitoring of cancer.

ige 232; 299pp; English.

invention relates to the isolation of novel cDNA sequences or at least an immunogenic portion of human colon tumour he sequences of the invention are useful in pharmaceutical is and vaccines for the prevention and treatment of cancers ion cancer. They are also useful for the diagnosis and of such cancers. Antibodies to the colon tumour proteins and esenting cells that express polynucleotides encoding colon teins can be used to inhibit the development of cancers. T- react specifically with colon tumour proteins are useful for umour cells from samples (e.g. blood) and for cancer treatment. cleotides sequences are also useful in gene therapy. AAS57325- epresent the cDNA sequences of the invention that encode for f human colon tumour proteins

26 BP; 88 A; 100 C; 78 G; 60 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 4; Length 326;
milarity 72.7%; Pred. No. 9.6e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    | |    ||  | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97
```

tandard; cDNA; 326 BP.

2 (first entry)

encoding portion of a human colon tumour protein.

on tumour protein; colon cancer; gene therapy; cytostatic; ss.

ns.

7-A2.

01.

01; 2001WO-US009246.

00; 2000US-0191597P.
00; 2000US-0202024P.
00; 2000US-0202189P.

ORIXA CORP.

J, Xu J, King GE;

-611627/70.

XX
PT New colon tumor proteins and related nucleic acid, useful for t
PT prevention, diagnosis and monitoring of cancer.
XX
PS Claim 4; Page 275; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA seq
CC encoding for at least an immunogenic portion of human colon tum
CC proteins. The sequences of the invention are useful in pharmace
CC compositions and vaccines for the prevention and treatment of c
CC such as colon cancer. They are also useful for the diagnosis ar
CC monitoring of such cancers. Antibodies to the colon tumour prot
CC antigen presenting cells that express polynucleotides encoding
CC tumour proteins can be used to inhibit the development of cance
CC cells that react specifically with colon tumour proteins are us
CC removing tumour cells from samples (e.g. blood) and for cancer
CC The polynucleotides sequences are also useful in gene therapy.
CC AAS58880 represent the cDNA sequences of the invention that enc
CC portions of human colon tumour proteins
XX
SQ Sequence 326 BP; 88 A; 100 C; 78 G; 60 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 326;
Best Local Similarity 72.7%; Pred. No. 9.6e-06;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||     |    | |    ||  | | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97
```

RESULT 16
AAS58211
ID AAS58211 standard; cDNA; 326 BP.
XX
AC AAS58211;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #887 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytos
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US009246.
XX
PR 24-MAR-2000; 2000US-0191597P.
PR 04-MAY-2000; 2000US-0202024P.
PR 05-MAY-2000; 2000US-0202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
DR WPI; 2001-611627/70.
XX
PT New colon tumor proteins and related nucleic acid, useful for
PT prevention, diagnosis and monitoring of cancer.
XX
PS Claim 4; Page 197; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA s
CC encoding for at least an immunogenic portion of human colon t
CC proteins. The sequences of the invention are useful in pharma

ons and vaccines for the prevention and treatment of cancers
olon cancer. They are also useful for the diagnosis and
of such cancers. Antibodies to the colon tumour proteins and
esenting cells that express polynucleotides encoding colon
teins can be used to inhibit the development of cancers. T-
react specifically with colon tumour proteins are useful for
umour cells from samples (e.g. blood) and for cancer treatment.
cleotides sequences are also useful in gene therapy. AAS57325-
epresent the cDNA sequences of the invention that encode for
f human colon tumour proteins

26 BP; 88 A; 100 C; 78 G; 60 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 4; Length 326;
milarity 72.7%; Pred. No. 9.6e-06;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97

tandard; cDNA; 532 BP.

2 (first entry)

n cancer related cDNA SEQ ID NO 780.

on; cancer; cytostatic; tumour; gene therapy; vaccine; gene;

ns.

4-A2.

2.

1; 2001WO-US043704.

0; 2000US-0252222P.
1; 2001US-0267011P.
1; 2001US-0279670P.
1; 2001US-0304037P.

RIXA CORP.

Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;

608400/65.

ed tumor colon polynucleotide and polypeptide, useful for the
prevention and/or treatment of cancer, in particular colon

EQ ID NO 780; 266pp + Sequence Listing; English.

ion relates to a human colon tumour expressed polynucleotide
ng a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
defined nucleotide sequences (ABV8669-ABV89289); (ii)
s of (i); (iii) at least 20 contiguous residues of (i); (iv)
that hybridize to (i), under moderately stringent conditions;
ces having at least 75% or 90% identity to (i); or (vi)
variants of (i). The compositions and methods of the present
are useful for the diagnosis, prevention and/or treatment of
rticularly colon cancer. (I) can be used in gene therapy and

CC (I) and (II) are useful in pharmaceutical compositions such a
CC Note: The sequence data for this patent did not form part of
CC specification, but was obtained in electronic format directly
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 532 BP; 158 A; 131 C; 137 G; 106 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 532;
Best Local Similarity 72.7%; Pred. No. 1.1e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| | | | || | | ||
Db 21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
||||||||||||||||
Db 81 CCGTGGAGGACAAGGAT 97

RESULT 18
AAS58049
ID AAS58049 standard; cDNA; 571 BP.
XX
AC AAS58049;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #725 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytos
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US009246.
XX
PR 24-MAR-2000; 2000US-0191597P.
PR 04-MAY-2000; 2000US-0202024P.
PR 05-MAY-2000; 2000US-0202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
DR WPI; 2001-611627/70.
XX
PT New colon tumor proteins and related nucleic acid, useful for
PT prevention, diagnosis and monitoring of cancer.
XX
PS Claim 4; Page 172; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA se
CC encoding for at least an immunogenic portion of human colon tu
CC proteins. The sequences of the invention are useful in pharmac
CC compositions and vaccines for the prevention and treatment of
CC such as colon cancer. They are also useful for the diagnosis a
CC monitoring of such cancers. Antibodies to the colon tumour pro
CC antigen presenting cells that express polynucleotides encoding
CC tumour proteins can be used to inhibit the development of canc
CC cells that react specifically with colon tumour proteins are u
CC removing tumour cells from samples (e.g. blood) and for cancer
CC The polynucleotides sequences are also useful in gene therapy.
CC AAS58880 represent the cDNA sequences of the invention that en
CC portions of human colon tumour proteins
XX
SQ Sequence 571 BP; 142 A; 163 C; 125 G; 131 T; 0 U; 10 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 571;
Best Local Similarity 72.7%; Pred. No. 1.1e-05;

; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 97

standard; cDNA; 572 BP.

02 (first entry)

4 encoding portion of a human colon tumour protein.

lon tumour protein; colon cancer; gene therapy; cytostatic; ss.

ens.

27-A2.

01.

01; 2001WO-US009246.

00; 2000US-0191597P.
00; 2000US-0202024P.
00; 2000US-0202189P.

ORIXA CORP.

J, Xu J, King GE;

-611627/70.

tumor proteins and related nucleic acid, useful for treatment,
1, diagnosis and monitoring of cancer.

Page 234; 299pp; English.

: invention relates to the isolation of novel cDNA sequences
for at least an immunogenic portion of human colon tumour
The sequences of the invention are useful in pharmaceutical
ons and vaccines for the prevention and treatment of cancers
olon cancer. They are also useful for the diagnosis and
g of such cancers. Antibodies to the colon tumour proteins and
resenting cells that express polynucleotides encoding colon
oteins can be used to inhibit the development of cancers. T-
: react specifically with colon tumour proteins are useful for
tumour cells from samples (e.g. blood) and for cancer treatment.
ucleotides sequences are also useful in gene therapy. AAS57325-
epresent the cDNA sequences of the invention that encode for
of human colon tumour proteins

572 BP; 141 A; 168 C; 113 G; 125 T; 0 U; 25 Other;

56.4%; Score 43.4; DB 4; Length 572;
imilarity 72.7%; Pred. No. 1.1e-05;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 97

RESULT 20
ABK86206
ID ABK86206 standard; cDNA; 2019 BP.
XX
AC ABK86206;
XX
DT 24-SEP-2002 (first entry)
XX
DE cDNA encoding human carcino-embryonic antigen (CEA) variant.
XX
KW Human; carcino-embryonic antigen; CEA; cytostatic; antiviral;
KW immunostimulant; cell-mediated immune response; tumour; breas
KW virus infection; prostate cancer; colorectal cancer; pancreat
KW lymphoma; leukaemia; hepadnavirus; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; ge
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7. .2019
FT /*tag= a
FT /product= "Carcino-embryonic antigen (CEA) va
FT the signal sequence deleted"
XX
PN WO200240059-A2.
XX
PD 23-MAY-2002.
XX
PF 01-NOV-2001; 2001WO-US045626.
XX
PR 01-NOV-2000; 2000US-00704232.
XX
PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
PA (MINC/) MINCHEFF M S.
PA (LOUK/) LOUKINOV D I.
PA (ZOUB/) ZOUBAK S.
XX
PI Mincheff MS, Loukinov DI, Zoubak S;
XX
DR WPI; 2002-527524/56.
DR P-PSDB; AAU98922.
XX
PT Inducing a cell-mediated immune response against a target ant
PT reducing undesired cells and stimulating presentation of an a
PT cell, comprises administering a polynucleotide encoding a var
PT antigen.
XX
PS Disclosure; Page 123-126; 146pp; English.
XX
CC The invention relates to a method of inducing a cell-mediated
CC response against a cell comprising a target antigen (I) in a
CC treating a subject having undesired cells, for example tumour
CC virally infected cells (C), reducing the number of (C) in a su
CC stimulating presentation of (I) by a cell. This is done by ad
CC a polynucleotide (II) encoding a variant of (I), so that (II)
CC in a cell and cell-mediated immune response is induced. The m
CC used to treat prostate cancer, breast cancer, colorectal canc
CC pancreatic cancer, as well as lymphomas and leukaemias. The m
CC also useful in treating chronic viral infections such as thos
CC hepadnaviruses, lentiviruses (including human immunodeficienc
CC (HIV)), herpesviruses and the flaviviruses and pestiviruses.
CC sequence represents the coding sequence of human carcino-embr
CC antigen (CEA) variant, which has the signal sequence deleted,
CC target antigen in the method of the invention
XX
SQ Sequence 2019 BP; 546 A; 621 C; 425 G; 427 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2019;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0,

```
AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |     |  |     ||  |  | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1444

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1461
```

```
standard; DNA; 2031 BP.

03  (revised)
05  (first entry)

ryonic antigen DNA.

ryonic antigen; CEA; neoplastic diseases; ds.

ns.

      Location/Qualifiers
le    103. .2028
      /*tag=  a

1.

4.

4;   94EP-00103986.

3;   93EP-00810214.

FFMANN LA ROCHE & CO AG F.

,  Pelegrin A,  Mach J;

304461/38.
R60619.

ryonic antigen (CEA) derivs - useful as reagents in immunoassay
sis of neoplastic diseases.

age 18; 30pp; English.

s the DNA sequence which encodes carcinoembryonic antigen (CEA)
CEA is free from cross-reactive CEA-like antigens, it is
lly indistinguishable from the solution form of CEA shed from
ls, and it is devoid of ethanolamine. AAR60619 can be used in a
mposition for detecting neoplastic diseases in biological
r in an immunoassay process where it can specifically detect
ce of tumour cells in a biological sample e.g. blood. (Updated
2003 to correct PN field.)

031 BP; 551 A; 642 C; 417 G; 421 T; 0 U; 0 Other;

          56.4%;  Score 43.4;  DB 2;  Length 2031;
milarity   72.7%;  Pred. No. 1.6e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |     |  |     ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1534

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1551
```

```
AAT36495
ID   AAT36495 standard; cDNA; 2059 BP.
XX
AC   AAT36495;
XX
DT   16-OCT-2003  (revised)
DT   27-AUG-2003  (revised)
DT   13-NOV-1996  (first entry)
XX
DE   Immunogenic carcinoembryonic antigen cDNA.
XX
KW   Carcinoembryonic antigen; immunogen; breast cancer; lung canc
KW   colon cancer; therapy; immunotherapy; vaccine; baculovirus; v
KW   Spodoptera frugiperda; insect; pA9080 AcNPV-CEA; ss.
XX
OS   Autographa californica nucleopolyhedrovirus.
OS   Homo sapiens.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .7
FT                   /*tag=  a
FT                   /function= "7-adenine motif in the polyhedrin
FT                   leader"
FT   CDS             18. .2009
FT                   /*tag=  b
FT   sig_peptide     18. .71
FT                   /*tag=  c
FT                   /function= "AcNPV 61k protein signal sequence'
FT   mat_peptide     72. .2006
FT                   /*tag=  d
FT   misc_feature    72. .80
FT                   /*tag=  e
FT                   /function= "3 N-terminal residues of the bacul
FT                   (MGS12) not present in human CEA"
FT   misc_feature    81. .2006
FT                   /*tag=  f
FT                   /function= "mature CEA"
FT   primer_bind     complement(81. .103)
FT                   /*tag=  g
FT                   /note= "5' primer for CEA amplification"
FT   primer_bind     1985. .2006
FT                   /*tag=  g
FT                   /note= "3' primer for CEA amplification"
FT   terminator      2030. .2040
FT                   /*tag=  i
FT                   /function= "universal translation termination
FT                   contg. stop codons in all 3 reading frames in
FT                   MGS12"
XX
PN   WO9532286-A2.
XX
PD   30-NOV-1995.
XX
PF   19-MAY-1995;   95WO-US006373.
XX
PR   20-MAY-1994;   94US-00246981.
XX
PA   (MICR-) MICROGENESYS INC.
XX
PI   Smith G,  Volvovitz F,  Hackett C;
XX
DR   WPI; 1996-020581/02.
DR   P-PSDB; AAR98519.
XX
PT   Immunogenic carcinoembryonic antigen produced using insect cel
PT   baculovirus expression system - useful in cancer therapy.
XX
PS   Disclosure; Page 47-49; 61pp; English.
XX
CC   A portion (AAT36494) of pA9080 AcNPV-CEA vector codes for reco
CC   soluble, immunogenic carcinoembryonic antigen (rCEA) (AAR98519
CC   human CEA gene was cloned from colon adenocarcinoma LS174T (AT
```

PCR (see also AAT36493-94) and modified for expression in insect
replacement of the signal sequence with a 61 kDa baculovirus
equence, and deletion of DNA encoding the hydrophobic C-terminal
he construct was inserted into baculovirus vector pMGS12 contg.
irin promoter, giving plasmid pA9080 AcNPV-CEA. rCEA was produced
evels in Sf900+ insect cells grown in serum-free media and
to a purity of over 95%. It can be used as an immunogen in
protect against breast, lung or colon cancer. (Updated on 27-
to correct OS field.) (Updated on 16-OCT-2003 to standardise OS

2059 BP; 572 A; 623 C; 425 G; 439 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 2; Length 2059;
imilarity 72.7%; Pred. No. 1.6e-05;
; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1512

CCGTGGAGGACAAGGAT 77
|||||||||||||||||
CCGTGGAGGACAAGGAT 1529

standard; cDNA; 2097 BP.

96 (first entry)

oryonic antigen cDNA.

mplify; polymerse chain reaction; PCR; human;
lycoprotein; BGP; membrane-bound; carcinoembryonic antigen; CEA;
protein; PR1A3 epitope; anti-PR1A3 antibody;
l carcinoma; monoclonal antibody; ss.

ens.

Location/Qualifiers
le 1. .90
/*tag= a
le 91. .2094
/*tag= b

-A1.

)5.

)4; 94WO-GB001816.

)3; 93GB-00017423.

IPERIAL CANCER RES TECHNOLOGY.

Durbin H, Snary D, Stewart LM, Young S, Bates PA;

106813/14.
R65168.

les which bind carcinoembryonic antigen - used for the
and treatment of colorectal carcinoma and for isolation and
ons.

; Fig 12; 67pp; English.

ces given in AAQ82805-07 encode human biliary glycoprotein
on-specific cross-reacting antigen (NCA) and human membrane-
inoembryonic antigen (CEA) respectively. BGP and CAE may be

CC used in the prearation of chimaeric proteins which may be us
CC identify the PR1A3 antigen and to raise the antibody of the
CC The preferred chimaeras are given in AAR77435-38. NCA may be
CC identify regions of importance in the CEA protein by compari
CC on the amino acid side chains with that in the binding regio
CC PR1A3 antibody (see also AAQ82785-90). The PR1A3 epitope was
CC an epitope of CEA within the B3-GPI region. Chimaeras in whi
CC was not membrane bound did not react with anti-PR1A3 antibod
CC indicating that the PR1A3 epitope is not present in non- mem
CC hybrid proteins. Antibodies which recognise the PR1A3 epitop
CC the detection of well and poorly differentiated colorectal c
CC The isolation of the specific PR1A3 epitope allows the develo
CC monoclonal antibodies specific for colorectal carcinoma. The
CC in the study, isolation and purification of molecules to whi
CC specifically bind and the imaging and treatment of cells exhi
CC molecules
XX
SQ Sequence 2097 BP; 559 A; 656 C; 439 G; 443 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 2; Length 2097;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
|||||||||||||||||||||| | | | || | | ||
Db 1463 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1523 CCGTGGAGGACAAGGAT 1539

RESULT 24
AAI72497
ID AAI72497 standard; DNA; 2105 BP.
XX
AC AAI72497;
XX
DT 16-MAY-2002 (first entry)
XX
DE CEA agonist coding sequence #2.
XX
KW CAP-1; epitope; carcinoembryonic antigen; CEA; agonist; immun
KW carcinoma; gastrointestinal; breast; pancreatic; bladder; ova
KW prostate; cancer; gene; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1. .2105
FT /*tag= a
FT /product= "CEA agonist polypeptide"
XX
PN WO200210379-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-CA001092.
XX
PR 31-JUL-2000; 2000US-0222043P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
PA (THER-) THERION BIOLOGICS.
PA (USSH) US NAT CANCER INST.
XX
PI Berinstein N, Tartaglia J, Tine JA, Panicali DL, Gritz L;
PI Schlom J;
XX
DR WPI; 2002-206189/26.
XX
PT Carcinoembryonic antigen agonist polypeptide for inducing an i
PT response in animal against antigen and for inhibiting an epito
PT expressing carcinoma cell, comprises a modified antigen epitop

'age 66-67; 69pp; English.

:nce encodes the carcinoembryonic antigen (CEA) agonist
le of the invention. This sequence represents the sequence given
[ID listing in the specification, and does not directly encodes
gonist polypeptide given in AAB47919. The CEA agonist contains a
'AP-1 epitope of CEA, in which position 6 is modified from Asp
increase its immunogenicity. The CEA agonist polypeptide of the
or DNA encoding it, are useful for: (i) inducing an immune
n an animal directed against a CEA protein or fragment, CEA
CEA epitope, a modified CEA epitope, cells expressing or
CEA protein or fragment; and (ii) inhibiting a CEA epitope
carcinoma cell, which is a gastrointestinal, breast,
, bladder, ovarian, lung or prostate carcinoma cell in a
ence is useful for manufacture of a medicament for the
of cancer

105 BP; 555 A; 658 C; 441 G; 451 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 6; Length 2105;
milarity 72.7%; Pred. No. 1.6e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1530

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1547

:andard; cDNA; 2106 BP.

: (first entry)

ircinoembryonic antigen (CEA) encoding cDNA SEQ ID NO:111.

iovirus; poxvirus; alphavirus; immune response; gp100;
gen; CEA; carcinoembryonic antigen; immunostimulant;
immunotherapy; interferon-gamma; IFN-gamma; cancer; ss.

d.

-A1.

.

; 2000WO-CA001253.

; 99US-0160879P.
; 2000US-0223325P.

NTIS PASTEUR LTD.

N, Tartaglia J, Moingeon P, Barber B;

08587/32.
97817.

mune response to tumor antigen, useful in immunotherapy of
administering the antigen to a lymphatic site.

Fig 8; 60pp; English.

invention describes a method for inducing an immune response,
l, to a tumour antigen (Ag) comprising administering Ag, or
d (I) that encodes it, to a lymphatic site. Cynomolgus monkeys

CC (Macaca fascicularis) were injected with a modified form of g
CC (a) into the left inguinal lymph node or (b) subcutaneously.
CC of (a) developed a cell-mediated response (indicated by produ
CC interferon-gamma from T lymphocytes when exposed to gp100 pep
CC only 2 of 4 animals of (b) did so. Also animals in (a) produc
CC greater antibody response to gp100. The method is used in imm
CC of a wide range of cancers through induction of a specific im
CC response (humoral and cellular) against the tumour antigens.
CC administered to a lymphatic site, Ag (or (I)) induces a stron
CC response than administration by other routes and may also bre
CC to Ag. AAB97708 and AAB97709 represent gp100 epitopes; AAB977
CC AAB97815 represent peptides derived from gp100 which stimulat
CC (IFN)-gamma production; AAH20120 encodes the modified gp100 p
CC in AAB97816; AAH20121 encodes the modified carcinoembryonic a
CC protein given in AAB97817; and AAB97818 represents a CEA modif
CC peptide, all of which are used in the exemplification of the
CC invention
XX
SQ Sequence 2106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 2106;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | | || | | |||
Db 1472 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1532 CCGTGGAGGACAAGGAT 1548

RESULT 26
AAD07347
ID AAD07347 standard; DNA; 2106 BP.
XX
AC AAD07347;
XX
DT 18-SEP-2001 (first entry)
XX
DE Modified carcinoembryonic antigen (CEA) DNA.
XX
KW Tumour-associated antigen; TAA; cytostatic; vaccine; gene ther
KW immune response; tetanus toxoid; TT; diphtheria toxoid; DT; pr
KW cancer; therapeutic; carcinoembryonic antigen; CEA; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1. .2106
FT /*tag= a
FT /product= "Modified carcinoembryonic antigen (C
XX
PN WO200149317-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-CA000005.
XX
PR 05-JAN-2000; 2000US-0174587P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Emtage P, Barber BH, Sambhara S, Sia CDY;
XX
DR WPI; 2001-441790/47.
DR P-PSDB; AAE05117.
XX
PT Enhancing immune response to antigen such as tumor antigen for
PT cancer in an animal involves administering an inducing agent to
PT animal followed by administering inducing agent-antigen mixture
XX

ig 3; 62pp; English.

ion relates to a method of enhancing an immune response against
ociated antigens (TAAs), such as GP100 and carcinoembryonic
EA) in an animal. The method involves priming of the animal
ducing agent such as tetanus toxoid (TT) or diphtheria toxoid
equently followed by administration of an inducing agent-
xture. The method provides the enhancement or augmentation of
response to the antigen and/or improves a vaccination protocol
g use of less antigen. The immunisation of the animal with
ociated antigen is useful for the prophylactic or therapeutic
of cancer. The present DNA sequence encodes modified
ryonic antigen (CEA) related to the invention

106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 5; Length 2106;
milarity 72.7%; Pred. No. 1.6e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1531

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1548

tandard; DNA; 2106 BP.

2 (first entry)

t coding sequence #1.

tope; carcinoembryonic antigen; CEA; agonist; immune response;
gastrointestinal; breast; pancreatic; bladder; ovarian; lung;
cancer; gene; ds.

Location/Qualifiers
1. .2106
/*tag= a
/product= "CEA agonist polypeptide"

9-A2.

2.

1; 2001WO-CA001092.

0; 2000US-0222043P.

ENTIS PASTEUR LTD.
ERION BIOLOGICS.
NAT CANCER INST.

N, Tartaglia J, Tine JA, Panicali DL, Gritz L;

-206189/26.
AB47918.

ryonic antigen agonist polypeptide for inducing an immune
in animal against antigen and for inhibiting an epitope antigen
carcinoma cell, comprises a modified antigen epitope.

Fig 1; 69pp; English.

XX
CC This sequence encodes the carcinoembryonic antigen (CEA) agoni
CC polypeptide of the invention. This sequence represents the seq
CC in the figures in the specification, and it directly encodes t
CC agonist polypeptide given in AAB47918. The CEA agonist contain
CC modified CAP-1 epitope of CEA, in which position 6 is modified
CC to Asn to increase its immunogenicity. The CEA agonist polypep
CC invention, or DNA encoding it, are useful for: (i) inducing an
CC response in an animal directed against a CEA protein or fragme
CC agonist, a CEA epitope, a modified CEA epitope, cells expressi
CC binding a CEA protein or fragment; and (ii) inhibiting a CEA e
CC expressing carcinoma cell, which is a gastrointestinal, breast
CC pancreatic, bladder, ovarian, lung or prostate carcinoma cell
CC patient, hence is useful for manufacture of a medicament for t
CC treatment of cancer
XX
SQ Sequence 2106 BP; 559 A; 658 C; 442 G; 447 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2106;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||||||||||||||||||||| | | | || | | |||
Db 1472 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1532 CCGTGGAGGACAAGGAT 1548

RESULT 28
ADE13860
ID ADE13860 standard; DNA; 2106 BP.
XX
AC ADE13860;
XX
DT 29-JAN-2004 (first entry)
XX
DE CEA-CAP6D nucleotide sequence SEQ ID NO:23.
XX
KW carcinoembryonic antigen; CEA; CEA(6D)-1; 2; cytostatic; vacci
KW tumour antigen; immunotherapy; gene; ds.
XX
OS Unidentified.
XX
PN WO2003085087-A2.
XX
PD 16-OCT-2003.
XX
PF 09-APR-2003; 2003WO-US010916.
XX
PR 09-APR-2002; 2002US-0372972P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
PA (THER-) THERION BIOLOGICS INC.
XX
PI Parrington M, Zhang L, Rovinski B, Gritz LR, Greenhalgh T;
XX
DR WPI; 2003-877029/81.
XX
PT New isolated DNA molecule comprising the carcinoembryonic anti
PT 1,2 sequence, useful for diagnosing, preventing and treating c
PT determining the effectiveness of a chemotherapeutic or other t
PT regimen.
XX
PS Example 1; SEQ ID NO 23; 56pp; English.
XX
CC The present invention describes an isolated DNA molecule compr
CC carcinoembryonic antigen (CEA) (6D)-1,2 sequence of 2106 bp (s
CC ADE13861), or its fragment. Also described: (1) an expression
CC comprising the nucleic acid sequence CEA(6D)-1,2, or its fragr
CC describes above; (2) a composition comprising the expression

)harmaceutical carrier; and (3) preventing or treating cancer
; administering to a host the expression vector of (1). CEA(6D)-
'tostatic activity, and can be used in vaccines. The CEA(6D)-1,2
:id and target polypeptide are useful for diagnosing, preventing
.ng cancer, predicting prognosis, or determining the
iess of a chemotherapeutic or other treatment regimen. The
1 vector may be used for the insertion and expression of CEA(6D)
:ic acid encoding tumour antigens for the immunotherapeutic
of cancer. The target polypeptides are useful in generating
: used in screening assays or for immunotherapy. The present
epresents the CEA-CAP6D nucleotide sequence, which is given in
: with CEA(6D)-1,2 in the exemplification of the present

106 BP; 559 A; 659 C; 442 G; 446 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 9; Length 2106;
milarity 72.7%; Pred. No. 1.6e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1531

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1548

tandard; cDNA; 2109 BP.

2 (first entry)

:yonic antigen coding sequence.

:yonic antigen; CEA; epitope; human leucocyte antigen; HLA;
)compatibility complex; MHC; CD4; T lymphocyte; cancer; tumour;
antitumour; therapy; diagnosis; human; human; gene; ss.

is.

Location/Qualifiers
1. .2109
/*tag= a
/product= "CEA"

:-A2.

:.

; 2001WO-US028467.

; 2000US-0232185P.

'OS INC.

'omlinson A;

62345/39.
:75931.

human leukocyte antigen (HLA) class II-restricted epitopes,
solating a major histocompatibility complex (MHC) molecule
peptide and eluting the peptide.

Fig 1A; 98pp; English.

sequence is the coding sequence for human carcinoembryonic
A). Epstein-Barr virus-transformed B cell lines expressing CEA

CC encoding cDNA were used to identify human leucocyte antigen (
CC II binding CEA peptide epitopes. The invention provides metho
CC identifying peptide epitopes that activate CD4+ T lymphocyte
CC involved in the initiation, promotion, or exacerbation of cer
CC diseases. It is based on the discovery that it is possible to
CC human leucocyte antigen class II restricted epitopes naturall
CC by antigen-presenting cells transfected with DNA encoding a p
CC which the epitopes are derived. The method has 2 phases: immu
CC mass fingerprinting, and epitope verification. By applying th
CC the tumour antigen CEA, CEA-derived peptides (see ABB75917-26
CC identified as epitopes that could be involved in the pathogen
CC cancer in human patients expressing the DR4 or DR1 class II m
CC histocompatibility complex (MHC) allele. The epitopes, and alt
CC peptide ligands (APL), are used in claimed methods of activati
CC reactivity, altering T cell response, diagnosis, and identifyi
CC reagent for diagnosing cancer. Also claimed are methods of car
CC therapy, and of identifying a class II MHC-binding fragment of
CC antigen. APLs are also useful as vaccines
XX
SQ Sequence 2109 BP; 561 A; 659 C; 442 G; 447 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2109;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||||||||||||||||||||| | | | || | | |||
Db 1475 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1535 CCGTGGAGGACAAGGAT 1551

RESULT 30
AAT33302
ID AAT33302 standard; cDNA; 2220 BP.
XX
AC AAT33302;
XX
DT 12-NOV-1996 (first entry)
XX
DE Carcinoembryonic antigen gene.
XX
KW Poliovirus; encapsidation; capsid; vaccine; genetic immunisati
KW carcinoembryonic antigen; CEA; tumour; cancer therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .2206
FT /*tag= a
XX
PN WO9625173-A1.
XX
PD 22-AUG-1996.
XX
PF 13-FEB-1996; 96WO-US001895.
XX
PR 15-FEB-1995; 95US-00389459.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Morrow CD, Porter DC, Ansardi DC;
XX
DR WPI; 1996-393136/39.
DR P-PSDB; AAW00182.
XX
PT Encapsidation of recombinant polio:virus nucleic acid for use i
PT - using a polio:virus nucleic acid which lacks the P1 capsid re
PT an expression system which provides the region.
XX
PS Example 7; Page 68-71; 108pp; English.

quence (AAT33302) codes for the human carcinoembryonic antigen
000182). The CEA can be used to substitute the P1 capsid gene of
s in recombinant poliovirus nucleic acids (rPNAs). Such rPNAs
sidated by introduction into a host cell together with a
virus or plasmid vector encoding the poliovirus P1 capsid
protein. Encapsidated rPNAs are useful for genetic
.on, stimulating an immune response to CEA as a means of cancer

2220 BP; 596 A; 693 C; 470 G; 461 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 2; Length 2220;
.milarity 72.7%; Pred. No. 1.6e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||| | | | || | | |||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1630

CGTGGAGGACAAGGAT 77
||||||||||||||||
CGTGGAGGACAAGGAT 1647

standard; DNA; 2349 BP.

03 (revised)
05 (first entry)

pression cassette from pH6.CEA.HA.

chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
; plasmid; pGEM.CEA; vaccinia; H6 promoter; amplify; primer;
scription termination signal; T5NT; FCR3 strain; ALVAC C3;
rum; 42K entomopox promoter; VQCP3L; donor plasmid; pSD414;
olymerase chain reaction; poxvirus; antigenic response;
.cal response; pathogen; ss.

Location/Qualifiers
re 1. .59
/*tag= a
/note= "Flanking sequence"
60. .183
/*tag= b
/note= "Vaccinia H6 promoter"
184. .2292
/*tag= c
/product= "CEA"
re 2293. .2349
/*tag= d
/note= "Flanking sequence"

-A1.

04.

04; 94WO-US000888.

03; 93US-00007115.
04; 94US-00184009.

IROGENETICS CORP.

, Tartaglia J, Cox WI;

-263767/32.

XX
PT Attenuated recombinant virus used for cancer therapy - compri
PT encoding cytokine and/or tumour associated antigen.
XX
PS Example 17; Fig 23; 232pp; English.
XX
CC The sequences given in AAQ67868-69 represent H6/CEA expressio
CC and flanking regions from plasmids pH6.CEA.C3.2 and pH6.CEA.H
CC respectively. These sequences were used in the construction o
CC ALVAC-based recombinant viruses containing the CEA gene. The
CC 5' and 3'untranslated regions (UTRs) are contained in plasmid
CC The 5' end of the CEA construct was modified to remove the 5'
CC place the vaccinia H6 promoter before the ATG initiation codo
CC primers CEA1 and CEA2. The resulting fragment links the 3' 30
CC H6 promoter to the CEA initiation codon and the next 28 bp of
CC gene. The 3' UTR was also removed and replaced by the vaccini
CC transcription termination signal (T5NT) followed by a series
CC restriction sites, using primers CEA3 and CEA4. The SERA cDNA
CC isolated from the FCR3 strain of P. falciparum and modified t
CC the T5NT sequence and a series of restriction sites, at the 3
CC 5' end was modified by replacing the 5' untranslated region (
CC several restriction sites and the 42K entomopox promoter. Thi
CC expression cassette was cloned into a BamHI/XhoI-digested VQC
CC fragment to generate an ALVAC C3 donor plasmid. The pSD553 va
CC plasmid was then constructed by PCR using pSD414 as a templat
CC may also be used in the production of the ALVAC C3 donor plas
CC gene sequence was introduced into ALVAC by recombination betw
CC linearised CEA.C3 donor plasmid and ALVAC rescuing plamid. Th
CC viruses may be used in a composition for inducing an antigenic
CC immunological response, ie. for immunisation against pathogens
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2349 BP; 640 A; 693 C; 477 G; 539 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 2; Length 2349;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
||||||||||||||||||||||| | | | || | | |||
Db 1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
||||||||||||||||
Db 1718 CCGTGGAGGACAAGGAT 1734

RESULT 32
AAZ08470
ID AAZ08470 standard; DNA; 2349 BP.
XX
AC AAZ08470;
XX
DT 19-OCT-1999 (first entry)
XX
DE H6/CEA expression cassette and flanking regions from pH6.CEA.H
XX
KW Attenuated recombinant virus; cytokine; tumour associated anti
KW NYVAC recombinant virus; ALVAC recombinant virus; gene therapy
KW cancer; tumour necrosis factor; nuclear phosphoprotein; p53; I
KW interleukin; interferon; IFN-gamma; IL-4; melanoma associated
KW carcinoembryonic antigen; immunisation; antigenic; poxvirus; i
KW immunological response; immunotherapy; vaccine; Newcastle Dise
XX
OS Synthetic.
OS Vaccinia virus.
XX
PN US5942235-A.
XX
PD 24-AUG-1999.
XX
PF 02-JUN-1995; 95US-00458356.
XX

)1; 81US-00334456.
)2; 82US-00446824.
)4; 84US-00622135.
)7; 87US-00090209.
)7; 87US-00090711.
)7; 87US-00110335.
)8; 88US-00186054.
)8; 88US-00234390.
)0; 90US-00537882.
)0; 90US-00537890.
)1; 91US-00805567.
)2; 92US-00847977.
)2; 92US-00847951.
)2; 92US-00881995.
2; 92US-00918278.
3; 93US-00007115.
4; 94US-00184009.
4; 94US-00228926.
4; 94US-00306259.

ALTH RES INC.

;

493494/41.

t poxviruses comprising exogenous DNA encoding antigenic
ts useful in immunotherapy to immunize against cancers and
ases such as influenza, Newcastle Disease and rabies.

; Fig 23; 163pp; English.

t invention describes a recombinant poxvirus (I), comprising
DNA encoding an antigenic determinant of a pathogen which is
ssed in vivo in infected host cells after administration to a
d therefore induces an immunological response. (I) may be used
te patients against a wide range of diseases and disorders
on the type of antigen encoded by the exogenous DNA. (I) may be
ccinate against diseases such as rabies, influenza and
Disease. It is particularly useful for immunising against
he poxvirus (I) also provides a means of manipulating
s and tumour cells for use in cell-based immunotherapeutic
for cancer. (I) also have enhanced safety compared to
ed viruses (attenuation reduces the virulence of the viruses)
recombinant poxvirus vaccines. This increased level of safety
e possibility of a 'runaway' infection in the host and reduces
of transmission from vaccinated to unvaccinated individuals
ination of the environment. The present sequence represents a
ression cassette and flanking regions from pH6.CEA.HA used in
ification of the present invention

349 BP; 640 A; 693 C; 477 G; 539 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 2; Length 2349;
ilarity 72.7%; Pred. No. 1.6e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | || | | |||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1717

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1734

tandard; DNA; 2389 BP.

(first entry)

XX
DE Human carcino-embryonal antigen encoding DNA.
XX
KW Human; MUC1; carcino-embryonal antigen; CEA; treatment; tumor
KW tumor antigen; anticancer; vaccine; immunotherapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .2389
FT /*tag= a
FT /product= "carcino-embryonal antigen"
FT /transl_exception= (pos:847. .849,aa:Asp)
FT /note= "This sequence contains in-frame stop
XX
PN WO200124832-A2.
XX
PD 12-APR-2001.
XX
PF 26-SEP-2000; 2000WO-DE003443.
XX
PR 27-SEP-1999; 99DE-01048105.
XX
PA (PECH/) PECHER G.
XX
PI Pecher G;
XX
DR WPI; 2001-266243/27.
DR P-PSDB; AAB70880.
XX
PT Composition for treating or preventing tumors, comprises plasm
PT optionally recombinant viruses, that express mucin and carcino
PT antigen.
XX
PS Claim 1; Fig 2; 11pp; German.
XX
CC This invention describes a novel composition (A) for treating
CC preventing human tumors that express the tumor antigens mucin
CC carcino-embryonal antigen (CEA). The composition has anticance
CC comprises (1) a plasmid containing the human mucin gene MUC1 (
CC active fragments or a sequence that expresses at least three r
CC specified sequence (1)); and/or (2) a plasmid containing the g
CC human CEA. (A) optionally used in combination with recombinant
CC that express the same tumor antigens, are useful as vaccines f
CC immunotherapy and prevention of tumors that express mucin and/
CC combination of antigens and vectors improves activity of the c
CC The method does not require manipulation of tumor cells and is
CC specific for a particular patient. The vaccines can be adminis
CC directly in vivo. This sequence encodes the human carcino-embr
CC antigen (CEA) protein used in the method of the invention
XX
SQ Sequence 2389 BP; 637 A; 730 C; 506 G; 516 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 4; Length 2389;
Best Local Similarity 72.7%; Pred. No. 1.6e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
||||||||||||||||||||| | | | || | | |||
Db 1475 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||
Db 1535 CCGTGGAGGACAAGGAT 1551

RESULT 34
AAQ67868
ID AAQ67868 standard; DNA; 2434 BP.
XX
AC AAQ67868;
XX
DT 25-MAR-2003 (revised)

'95 (first entry)

:pression cassette from pH6.CEA.C3.2.

:e chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
:A; plasmid; pGEM.CEA; vaccinia; H6 promoter; amplify; primer;
inscription termination signal; T5NT; FCR3 strain; ALVAC C3;
)arum; 42K entomopox promoter; VQCP3L; donor plasmid; pSD414;
)olymerase chain reaction; poxvirus; antigenic response;
;ical response; pathogen; ss.

:.

Location/Qualifiers
ure 1. .56
/*tag= a
/note= "Flanking sequence"
57. .180
/*tag= b
/note= "Vaccinia H6 promoter"
181. .2289
/*tag= c
/product= "CEA"
ure 2290. .2434
/*tag= d
/note= "Flanking sequence"

-A1.

94.

94; 94WO-US000888.

93; 93US-00007115.
94; 94US-00184009.

IROGENETICS CORP.

3, Tartaglia J, Cox WI;

-263767/32.

i recombinant virus used for cancer therapy - comprises DNA
:ytokine and/or tumour associated antigen.

7; Fig 22; 232pp; English.

ices given in AAQ67868-69 represent H6/CEA expression cassettes
ing regions from plasmids pH6.CEA.C3.2 and pH6.CEA.HA
ely. These sequences were used in the construction of NYVAC- and
:d recombinant viruses containing the CEA gene. The CEA gene and
intranslated regions (UTRs) are contained in plasmid pGEM.CEA.
l of the CEA construct was modified to remove the 5' UTR and
vaccinia H6 promoter before the ATG initiation codon using
:A1 and CEA2. The resulting fragment links the 3' 30 bp of the
:r to the CEA initiation codon and the next 28 bp of the CEA
3' UTR was also removed and replaced by the vaccinia early
:ion termination signal (T5NT) followed by a series of
)n sites, using primers CEA3 and CEA4. The SERA cDNA was
from the FCR3 strain of P. falciparum and modified to include
:equence and a series of restriction sites, at the 3' end. The
; modified by replacing the 5' untranslated region (UTR) with
:striction sites and the 42K entomopox promoter. This SERA
l cassette was cloned into a BamHI/XhoI-digested VQCP3L vector
:o generate an ALVAC C3 donor plasmid. The pSD553 vaccinia donor
is then constructed by PCR using pSD414 as a template. pSD553
)e used in the production of the ALVAC C3 donor plasmid. The CEA
:nce was introduced into ALVAC by recombination between NotI-
l CEA.C3 donor plasmid and ALVAC rescuing plamid. The resulting
y be used in a composition for inducing an antigenic or
.cal response, ie. for immunisation against pathogens. (Updated
2003 to correct PN field.)

SQ Sequence 2434 BP; 668 A; 704 C; 495 G; 567 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 2; Length 2434;
Best Local Similarity 72.7%; Pred. No. 1.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels (

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAA
Db 1655 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAA

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1715 CCGTGGAGGACAAGGAT 1731

RESULT 35
AAZ08469
ID AAZ08469 standard; DNA; 2434 BP.
XX
AC AAZ08469;
XX
DT 19-OCT-1999 (first entry)
XX
DE H6/CEA expression cassette and flanking regions from pH6.CEA.
XX
KW Attenuated recombinant virus; cytokine; tumour associated ant
KW NYVAC recombinant virus; ALVAC recombinant virus; gene therap
KW cancer; tumour necrosis factor; nuclear phosphoprotein; p53;
KW interleukin; interferon; IFN-gamma; IL-4; melanoma associated
KW carcinoembryonic antigen; immunisation; antigenic; poxvirus;
KW immunological response; immunotherapy; vaccine; Newcastle Dis
XX
OS Synthetic.
OS Vaccinia virus.
XX
PN US5942235-A.
XX
PD 24-AUG-1999.
XX
PF 02-JUN-1995; 95US-00458356.
XX
PR 24-DEC-1981; 81US-00334456.
PR 08-DEC-1982; 82US-00446824.
PR 19-JUN-1984; 84US-00622135.
PR 27-AUG-1987; 87US-00090209.
PR 28-AUG-1987; 87US-00090711.
PR 20-OCT-1987; 87US-00110335.
PR 25-APR-1988; 88US-00186054.
PR 23-AUG-1988; 88US-00234390.
PR 14-JUN-1990; 90US-00537882.
PR 14-JUN-1990; 90US-00537890.
PR 16-DEC-1991; 91US-00805567.
PR 03-MAR-1992; 92US-00847977.
PR 06-MAR-1992; 92US-00847951.
PR 04-MAY-1992; 92US-00881995.
PR 22-JUL-1992; 92US-00918278.
PR 20-JAN-1993; 93US-00007115.
PR 19-JAN-1994; 94US-00184009.
PR 14-APR-1994; 94US-00228926.
PR 13-SEP-1994; 94US-00306259.
XX
PA (HEAL-) HEALTH RES INC.
XX
PI Paoletti E;
XX
DR WPI; 1999-493494/41.
XX
PT Recombinant poxviruses comprising exogenous DNA encoding antig
PT determinants useful in immunotherapy to immunize against cance
PT other diseases such as influenza, Newcastle Disease and rabies
XX
PS Example 17; Fig 22; 163pp; English.
XX

it invention describes a recombinant poxvirus (I), comprising
DNA encoding an antigenic determinant of a pathogen which is
ssed in vivo in infected host cells after administration to a
id therefore induces an immunological response. (I) may be used
ite patients against a wide range of diseases and disorders
on the type of antigen encoded by the exogenous DNA. (I) may be
ccinate against diseases such as rabies, influenza and
Disease. It is particularly useful for immunising against
he poxvirus (I) also provides a means of manipulating
s and tumour cells for use in cell-based immunotherapeutic
for cancer. (I) also have enhanced safety compared to
ed viruses (attenuation reduces the virulence of the viruses)
recombinant poxvirus vaccines. This increased level of safety
e possibility of a 'runaway' infection in the host and reduces
of transmission from vaccinated to unvaccinated individuals
ination of the environment. The present sequence represents a
ression cassette and flanking regions from pH6.CEA.C3.2 used in
ification of the present invention

434 BP; 668 A; 704 C; 495 G; 567 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 2; Length 2434;
milarity 72.7%; Pred. No. 1.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||     |   |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1714

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1731
```

tandard; DNA; 2907 BP.

4 (first entry)

-12 coding sequence.

Hypertensive; Respiratory; Anti-HIV; Antiallergic;
ctive; Nootropic; Antianemic; Antiarteriosclerotic;
matory; Opthalmological; Muscular; Hepatotropic;
ctive; Antiasthmatic; Anticonvulsant; Virucide; Antibacterial;
Antiparasitic; Protozoacide; Antihelminthic; Cytostatic;
ective; Antiparkinsonian; Antipsoriatic; Antigout;
ic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;
l growth; cell differentiation; cell death; CGDD;
ferative disorder; cancer; developmental disorder;
al disorder; autoimmune disorder; inflammatory disorder;
reproductive disorder; gene; ds.

ıs.

'5-A2.

1.

1; 2003WO-US008310.

2; 2002US-0364494P.
2; 2002US-0369129P.
2; 2002US-0372511P.

YTE GENOMICS INC.

Tran UK, Hafalia AJA, Burford N, Honchell CD;
PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;
Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;

PI Becha SD, Lee SY, Swarnakar A, Yue H, Warren BA, Baughn
PI Lee S, Ho A, Gandhi AR, Yao MG;
XX
DR WPI; 2003-779081/73.
DR P-PSDB; ADD78231.
XX
PT New polypeptides and polynucleotides associated with cell gro
PT differentiation and death, useful for diagnosing, treating or
PT e.g. developmental, neurological, autoimmune, inflammatory or
PT reproductive disorders.
XX
PS Claim 5; SEQ ID NO 51; 320pp; English.
XX
CC The present invention relates to novel human proteins (I; ADD
CC ADD78258) and their coding sequences (II; ADD78259-ADD78297),
CC associated with cell growth, differentiation and death, refer
CC CGDD-n proteins, where n is a number from 1 to 39. The CGDD p
CC their coding sequences are useful for diagnosing, treating or
CC cell proliferative disorders (e.g. cirrhosis, hepatitis,
CC arteriosclerosis, psoriasis, primary thrombocytopenia) or can
CC adenocarcinoma, sarcoma or cancers of the bone, bone marrow,
CC breast, colon, kidney, liver, lung or uterus), developmental
CC (e.g. renal tubular acidosis, Becker muscular dystrophy, gona
CC dysgenesis, hypothyroidism or seizures), neurological disorde
CC Pick's disease, cataract, epilepsy, ischemic cerebrovascular
CC stroke, Alzheimer's disease, Parkinson's disease or dementia),
CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anemi
CC diabetes mellitus, bronchitis, osteoporosis, osteoarthritis,
CC arthritis, contact dermatitis or gout), viral, bacterial, fung
CC parasitic, protozoan or helminthic infections, reproductive di
CC (e.g. infertility, ectopic pregnancy, premature ovarian failur
CC puberty or prostatitis) or disorders of the placenta (e.g. pre
CC choriocarcinoma, placenta previa, placental or maternal floor
CC or chronic villitis).
XX
SQ Sequence 2907 BP; 782 A; 898 C; 611 G; 616 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 9; Length 2907;
Best Local Similarity 72.7%; Pred. No. 1.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||     |   |  |    ||  |  |||
Db       2106 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       2166 CCGTGGAGGACAAGGAT 2182
```

RESULT 37
AAN81611
ID AAN81611 standard; DNA; 2928 BP.
XX
AC AAN81611;
XX
DT 17-NOV-1990 (first entry)
XX
DE Carcinoembryonic antigen gene.
XX
KW Carcinoembryonic antigen; antibody; tumor diagnosis; ssDNA.
XX
OS Homo sapiens.
XX
PN JP63177794-A.
XX
PD 21-JUL-1988.
XX
PF 14-JAN-1987; 87JP-00006851.
XX
PR 14-JAN-1987; 87JP-00006851.
XX
PA (SUNR) SUNTORY LTD.

245625/35.
P95424.

inoembryonic antigen - used to produce antibodies and detect
sue without reacting to CEA-related antibody.

; Page ?; 15pp; Japanese.

carcinoembryonic antigen (CEA) gene is used to express CEA.
raised against human CEA are used to detect tumor tissue
cinoma). See also AAP95424

928 BP; 824 A; 824 C; 584 G; 696 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 1; Length 2928;
milarity 72.7%; Pred. No. 1.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |      |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1522

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1539
```

tandard; cDNA; 2928 BP.

3 (revised)
7 (first entry)

inoembryonic antigen cDNA.

cinoembryonic antigen; CEA; antibody; colon; cancer; ds.

ns.

Location/Qualifiers
1. .2097
/*tag= a
e 1. .90
/*tag= b
e 91. .2094
/*tag= c

-A.

7.

7; 96JP-00221770.

7; 87JP-00006851.

NTORY LTD.

328482/30.
W22844.

inoembryonic antigen protein - used to prepare anti-CEA
.

age 9-12; 14pp; Japanese.

t sequence encodes human carcinoembryonic antigen (CEA), which
d in the preparation anti-CEA antibodies (Ab). PolyA+ RNA was
from the colon tissue of a colon cancer patient, and a lambda
library prepared. Positive clones were separated by Ab

CC screening and analysed using various anti-CEA Ab. A cDNA libra
CC prepared by the Okayama-Bark method. The cDNA clone, lambda Kr
CC reacted with 4 anti-CEA Ab, was judged to encode human CEA. (U
CC 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 2928 BP; 825 A; 820 C; 588 G; 695 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 2; Length 2928;
Best Local Similarity 72.7%; Pred. No. 1.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

```
Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||     |      |  |    ||  |  | |||
Db       1463 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1523 CCGTGGAGGACAAGGAT 1539
```

RESULT 39
ABL64746
ID ABL64746 standard; DNA; 2974 BP.
XX
AC ABL64746;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3083.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyrc
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; canc
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adeno
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.

0; 2000US-0237173P.
0; 2000US-0237278P.
0; 2000US-0237294P.
0; 2000US-0237295P.
0; 2000US-0237316P.
0; 2000US-0237425P.
0; 2000US-0237598P.
0; 2000US-0237604P.
0; 2000US-0237606P.
0; 2000US-0237608P.
0; 2000US-0244867P.
0; 2000US-0245084P.

ALON PHARM.

Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Weaver Z;

188264/24.

for anti-neoplastic agent involves exposing cells to a chemical
e tested for anti-neoplastic activity, and determining a change
ion of a gene of a signature gene set.

EQ ID NO 3083; 44pp; English.

: invention describes a method (M1) for screening for an anti-
agent. The method involves exposing cells to a chemical agent
ed for anti-neoplastic activity, determining a change in
of at least one gene (I) of a signature gene set, where (I)
a sequence (S) selected from 8447 sequences (given in ABL61664
)), or is at least 95% identical to (S), where a change in
is indicative of anti-neoplastic activity. (I) has cytostatic
id can be used in gene therapy. M1 can be used for screening an
astic agent, and can be used for producing a product which is
ollected with respect to the anti-neoplastic agent as a result
the data is sufficient to convey the chemical structure and/or
of the agent. M1 can be used in the treatment of cancer such
reast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
: pancreatic cancer, adenocarcinoma, carcinoma, clear cell
iltrating ductal cancer, infiltrating lobular cancer, squamous
oma, neuroendocrine carcinoma, papillary carcinoma and Wilm's

74 BP; 840 A; 847 C; 613 G; 674 T; 0 U; 0 Other;

56.4%; Score 43.4; DB 6; Length 2974;
ilarity 72.7%; Pred. No. 1.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| | | | || | | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1665

andard; DNA; 2974 BP.

(first entry)

used to diagnose liver cancer.

cancer; ds; hepatocellular carcinoma; hepatotropic;
liver tumour; cytostatic; expression profile; disease state;
gression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepa
PT carcinoma or metastatic liver tumor in a patient, involves det
PT level of expression of two or more genes in a liver tissue sam
XX
PS Claim 1; SEQ ID NO 2317; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and det
CC progression of liver cancer, hepatocellular carcinoma or metas
CC tumour in a patient, and differentiating metastatic liver canc
CC hepatocellular carcinoma in a patient, involving detecting the
CC expression of two or more genes represented in ABN93503-ABN974
CC tissue sample. The method of the invention has hepatotropic, a
CC cytostatic activity. The method is useful for diagnosing and d
CC the progression of liver cancer, hepatocellular carcinoma and
CC liver carcinoma in a patient. The method is useful for identif
CC expression profiles which serve as useful diagnostic markers a
CC markers that can be used to monitor disease states, disease pr
CC drug toxicity, drug efficacy and drug metabolism. Note: The se
CC for this patent did not form part of the printed specification
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2974 BP; 840 A; 847 C; 613 G; 674 T; 0 U; 0 Other;

Query Match 56.4%; Score 43.4; DB 6; Length 2974;
Best Local Similarity 72.7%; Pred. No. 1.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||| | | | || | | |||
Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1649 CCGTGGAGGACAAGGAT 1665

Search completed: April 13, 2004, 18:34:52
Job time : 220.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic search, using sw model

April 13, 2004, 17:11:06 ; Search time 1107.5 Seconds
(without alignments)
3013.466 Million cell updates/sec

US-10-090-326-24
77
1 agacaatcacagtctctgcg.....aacccgtggaggacaaggat 77

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

3470272 seqs, 21671516995 residues

hits satisfying chosen parameters: 6940544

length: 0
length: 2000000000

: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

. is the number of results predicted by chance to have a score greater than or equal to the score of the result being and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Descrip
1	45	58.4	862	6	I08155	I08155
2	45	58.4	862	9	HUMCEALV	M17191
3	45	58.4	1623	6	AX497809	AX49780
4	45	58.4	2839	6	I08156	I08156
5	45	58.4	2839	6	I08166	I08166
6	43.4	56.4	111	6	AX260617	AX26061
7	43.4	56.4	190	6	AX260603	AX26060
8	43.4	56.4	326	6	AX261236	AX26123
9	43.4	56.4	326	6	AX261461	AX26146
10	43.4	56.4	326	6	AX261739	AX26173
11	43.4	56.4	571	6	AX261074	AX26107
12	43.4	56.4	572	6	AX261473	AX26147
13	43.4	56.4	1422	9	HUMCEAX	M16234
14	43.4	56.4	2019	6	AX505112	AX50511
15	43.4	56.4	2031	6	A39900	A39900
16	43.4	56.4	2031	6	I67748	I67748
17	43.4	56.4	2092	9	HSCEAASP	X16455
18	43.4	56.4	2097	6	A43169	A43169
19	43.4	56.4	2097	6	AR079553	AR07955
20	43.4	56.4	2106	6	AX133657	AX13365
21	43.4	56.4	2106	6	AX192349	AX19234
22	43.4	56.4	2106	6	AX393888	AX39388
23	43.4	56.4	2109	6	AX133977	AX13397
24	43.4	56.4	2109	6	AX468838	AX46883
25	43.4	56.4	2220	6	AR044683	AR04468
26	43.4	56.4	2349	6	AR052808	AR05280
27	43.4	56.4	2349	6	AR288121	AR28812
28	43.4	56.4	2434	6	AR052807	AR05280
29	43.4	56.4	2434	6	AR288120	AR28812
30	43.4	56.4	2888	9	BC034671	BC03467
31	43.4	56.4	2928	6	E01630	E01630
32	43.4	56.4	2928	6	E13123	E13123
33	43.4	56.4	2929	9	HUMCEA	M15042
34	43.4	56.4	2974	6	AX332574	AX33257
35	43.4	56.4	2974	6	AX409670	AX40967
36	43.4	56.4	2974	6	AX658324	AX65832
37	43.4	56.4	2974	6	AX677147	AX67714
38	43.4	56.4	2974	6	AX805532	AX80553
39	43.4	56.4	2974	9	HUMCEAF	M29540
40	43.4	56.4	3036	9	HUMANTCE	M17303
41	41	53.2	588	6	BD229772	BD22977
42	37.4	48.6	859	6	I08165	I08165
43	35	45.5	162661	2	AC146882	AC14688
44	35	45.5	193704	2	AC146845	AC14684
c 45	32.8	42.6	472	6	AR273307	AR27330

ALIGNMENTS

RESULT 1
I08155
LOCUS I08155 862 bp DNA linear PA
DEFINITION Sequence 1 from Patent EP 0346702.
ACCESSION I08155
VERSION I08155.1 GI:589126
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 862)
AUTHORS Barnett,T.R., Elting,J.J., Kamarck,M.E. and Kretschmer,
TITLE Transfectant cell lines which express isolated genes o
carcinoembryonic antigen family
JOURNAL Patent: EP 0346702-A2 1 20-DEC-1989;

```
      Location/Qualifiers
      1. .862
      /organism="unknown"
      /mol_type="unassigned DNA"


           58.4%;  Score 45;  DB 6;  Length 862;
nilarity   74.0%;  Pred. No. 6.1e-06;
  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| |   |     |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 591

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 608



CEALV                   862 bp    mRNA    linear   PRI 01-NOV-1994
in carcinoembronyic antigen mRNA, partial cds.
191
191.1  GI:180224
cinoembryonic antigen.
 sapiens (human)
 sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 862)
arck,M.E., Elting,J.J., Hart,J.T., Goebel,S.J., Rae,P.M.,
durft,M.A., Nedwin,J.J. and Barnett,T.R.
cinoembryonic antigen family: expression in a mouse L-cell
nsfectant and characterization of a partial cDNA in
eriophage lambda gt11
. Natl. Acad. Sci. U.S.A. 84 (15), 5350-5354 (1987)
60984
5415
ginal source text: Human carcinoma cell line LoVo (CCL229), cDNA
mRNA, clone lambda-cLV7.
      Location/Qualifiers
      1. .862
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="19q13.2"
      1. .862
      /gene="CEA"
      <1. .>862
      /gene="CEA"
      /note="carcinoembronyic antigen"
      /codon_start=3
      /protein_id="AAA51968.1"
      /db_xref="GI:180225"
      /db_xref="GDB:G00-119-054"
      /translation="GLHNHHPIKPFITSNNSNPVEDEDAVALTCEPEIQNTTYLWWVN
      NQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYECGIQNELSVDHSDPVILNVLYGPDD
      PTISPSYTYYRPGVNLSLSCHAASNPPAQYSWLIDGNIQQHTQELFISNITEKNSGLY
      TCQANNSASGHSRTTVKTITVSADVPKPSISSNNSKPVEDKDAVAFHCEPEAQNTTYL
      WWVNGQSLPVSPRLQLSNGNRTLTLFNVTRNDARAYVCGIQNSVSANRSDPVTLDVLY
      GPDTPIISPPP"


           58.4%;  Score 45;  DB 9;  Length 862;
milarity   74.0%;  Pred. No. 6.1e-06;
  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| |   |     |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 591

CGTGGAGGACAAGGAT 77
```

```
                ||||||||||||||||
Db        592 CCGTGGAGGACAAGGAT 608


RESULT 3
AX497809
LOCUS       AX497809                1623 bp    DNA     linear   PAT
DEFINITION  Sequence 14 from Patent WO0240673.
ACCESSION   AX497809
VERSION     AX497809.1  GI:23342948
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y., Liu,C. and
  TITLE     Compositions and methods relating to lung specific gene
            proteins
  JOURNAL   Patent: WO 0240673-A 14 23-MAY-2002;
            Diadexus, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1623
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           58.4%;  Score 45;  DB 6;  Length 1623;
  Best Local Similarity   74.0%;  Pred. No. 6.4e-06;
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||| |   |     |  |    ||  |  | |||
Db         51 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db        111 CCGTGGAGGACAAGGAT 127


RESULT 4
I08156
LOCUS       I08156                  2839 bp    DNA     linear   PAT
DEFINITION  Sequence 2 from Patent EP 0346702.
ACCESSION   I08156
VERSION     I08156.1  GI:589127
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2839)
  AUTHORS   Barnett,T.R., Elting,J.J., Kamarck,M.E. and Kretschmer,
  TITLE     Transfectant cell lines which express isolated genes of
            carcinoembryonic antigen family
  JOURNAL   Patent: EP 0346702-A2 2 20-DEC-1989;
FEATURES             Location/Qualifiers
     source          1. .2839
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           58.4%;  Score 45;  DB 6;  Length 2839;
  Best Local Similarity   74.0%;  Pred. No. 6.8e-06;
  Matches   57;  Conservative    0;  Mismatches   20;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||| |   |     |  |    ||  |  | |||
Db       1479 AGACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1539 CCGTGGAGGACAAGGAT 1555
```

```
166                    2839 bp    DNA     linear   PAT 02-DEC-1994
uence 8 from Patent EP 0346710.
166
166.1  GI:589121

nown.
nown.
lassified.
(bases 1 to 2839)
nett,T.R.D., Elting,J.J.D., Kamarck,M.E. and Kretschmer,A.D.
As coding for members of the carcinoembryonic antigen family
ent: EP 0346710-A2 8 20-DEC-1989;
      Location/Qualifiers
      1. .2839
      /organism="unknown"
      /mol_type="unassigned DNA"


          58.4%;  Score 45;  DB 6;  Length 2839;
milarity   74.0%;  Pred. No. 6.8e-06;
  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||| |   |     | |    || |  | ||||||||||
GACAATCACAGTCTCTGCGGACGTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1538

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1555



50617                  111 bp    DNA     linear   PAT 26-OCT-2001
uence 268 from Patent WO0173027.
50617
50617.1  GI:16509584

o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

gher,M.J., Xu,J. and King,G.E.
positions and methods for therapy and diagnosis of colon cancer
ent: WO 0173027-A 268 04-OCT-2001;
IXA CORPORATION (US)
      Location/Qualifiers
      1. .111
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"


          56.4%;  Score 43.4;  DB 6;  Length 111;
milarity   72.7%;  Pred. No. 2.1e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||      |     | |    || |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97
```

```
LOCUS       AX260603                 190 bp    DNA     linear   PAT
DEFINITION  Sequence 254 from Patent WO0173027.
ACCESSION   AX260603
VERSION     AX260603.1  GI:16509570
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Meagher,M.J., Xu,J. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of c
  JOURNAL   Patent: WO 0173027-A 254 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .190
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 190;
  Best Local Similarity  72.7%;  Pred. No. 2.2e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |     | |    || |  | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97


RESULT 8
AX261236
LOCUS       AX261236                 326 bp    DNA     linear   PAT
DEFINITION  Sequence 887 from Patent WO0173027.
ACCESSION   AX261236
VERSION     AX261236.1  GI:16510203
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Meagher,M.J., Xu,J. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of c
  JOURNAL   Patent: WO 0173027-A 887 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .326
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 326;
  Best Local Similarity  72.7%;  Pred. No. 2.3e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |     | |    || |  | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97


RESULT 9
AX261461
LOCUS       AX261461                 326 bp    DNA     linear   PAT
```

```
ience 1112 from Patent WO0173027.
:1461
:1461.1  GI:16510428

> sapiens (human)
> sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

gher,M.J., Xu,J. and King,G.E.
positions and methods for therapy and diagnosis of colon cancer
ent: WO 0173027-A 1112 04-OCT-2001;
IXA CORPORATION (US)
      Location/Qualifiers
      1. .326
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"


          56.4%;  Score 43.4;  DB 6;  Length 326;
milarity   72.7%;  Pred. No. 2.3e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||      |   |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97



51739                 326 bp    DNA     linear   PAT 26-OCT-2001
uence 1390 from Patent WO0173027.
61739
61739.1  GI:16510706

o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

gher,M.J., Xu,J. and King,G.E.
positions and methods for therapy and diagnosis of colon cancer
ent: WO 0173027-A 1390 04-OCT-2001;
IXA CORPORATION (US)
      Location/Qualifiers
      1. .326
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"


          56.4%;  Score 43.4;  DB 6;  Length 326;
milarity   72.7%;  Pred. No. 2.3e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||      |   |  |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 80

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 97



61074                 571 bp    DNA     linear   PAT 26-OCT-2001
uence 725 from Patent WO0173027.
```

```
ACCESSION   AX261074
VERSION     AX261074.1  GI:16510041
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Meagher,M.J., Xu,J. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of c
  JOURNAL   Patent: WO 0173027-A 725 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             56.4%;  Score 43.4;  DB 6;  Length 571;
  Best Local Similarity   72.7%;  Pred. No. 2.4e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |   |  |    ||  |  | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97


RESULT 12
AX261473
LOCUS       AX261473                572 bp    DNA     linear   PAT
DEFINITION  Sequence 1124 from Patent WO0173027.
ACCESSION   AX261473
VERSION     AX261473.1  GI:16510440
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Meagher,M.J., Xu,J. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of c
  JOURNAL   Patent: WO 0173027-A 1124 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .572
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             56.4%;  Score 43.4;  DB 6;  Length 572;
  Best Local Similarity   72.7%;  Pred. No. 2.4e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |   |  |    ||  |  | |||
Db         21 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||
Db         81 CCGTGGAGGACAAGGAT 97


RESULT 13
HUMCEAX
LOCUS       HUMCEAX                 1422 bp    mRNA    linear   PRI
DEFINITION  Human carcinoembryonic antigen mRNA, 3' end.
ACCESSION   M16234
```

```
34.1  GI:180240
repeat; carcinoembryonic antigen.
 sapiens (human)
 sapiens
ryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
bases 1 to 1422)
ermann,W., Ortlieb,B., Friedrich,R. and von Kleist,S.
ation and characterization of cDNA clones encoding the human
inoembryonic antigen reveal a highly conserved repeating
cture
. Natl. Acad. Sci. U.S.A. 84 (9), 2960-2964 (1987)
4247
671
inal source text: Human colon adenocarcinoma, cDNA to mRNA,
es pCEA[1-5].
t entry and printed copy of sequence [1] kindly provided by
mmermann, 01-JUL-1987.
     Location/Qualifiers
     1. .1422
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /map="19q13.2"
     1. .1422
     /gene="CEA"
     <1. .1422
     /gene="CEA"
     /product="carcinoembryonic antigen mRNA"
     <1. .1121
     /gene="CEA"
     /note="carcinoembryonic antigen"
     /codon_start=2
     /protein_id="AAA51972.1"
     /db_xref="GI:180241"
     /db_xref="GDB:G00-119-054"
     /translation="NSNPVEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSND
     NRTLTLLSVTRNDVGPYECGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVN
     LSLSCHAASNPPAQYSWLIDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTT
     VKTITVSAELPKPSISSNNSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQ
     LSNGNRTLTLFNVTRNDARAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYL
     SGANLNLSCHSASNPSPQYSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGR
     NNSIVKSITVSASGTSPGLSAGATVGIMIGVLVGVALI"
on   <1. .502
     /note="repeat copy A"
on   503. .1036
     /note="repeat copy B"
on   1241. .>1422
     /note="Alu repeat"
bp upstream of PstI site.

          56.4%;  Score 43.4;  DB 9;  Length 1422;
ilarity   72.7%;  Pred. No. 2.6e-05;
 Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||     |  |    ||  |  | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 545

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 562



5112                2019 bp    DNA     linear   PAT 27-SEP-2002
ence 5 from Patent WO0240059.
5112
5112.1  GI:23386419

 sapiens (human)
 sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
REFERENCE   1
  AUTHORS   Mincheff,M.S., Loukinov,D.I. and Zoubak,S.
  TITLE     Methods and compositions for inducing cell-mediated immu
            responses
  JOURNAL   Patent: WO 0240059-A 5 23-MAY-2002;
            American Foundation for Biological Research Inc. (US) ;
            Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serg
FEATURES             Location/Qualifiers
     source          1. .2019
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             7. .2019
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD48803.1"
                     /db_xref="GI:23386420"
                     /db_xref="REMTREMBL:CAD48803"
                     /translation="MAKLTIESTPFNVAEGKEVLLLVHNLPQHLFGY
                     RQIIGYVIGTQQATPGPAYSGREIIYPNASLLIQNIIQNDTGFYTLH
                     TGQFRVYPELPKPSISSNNSKPVEDKDAVAFTCEPETQDATYLWWVN
                     LSNGNRTLTLFNVTRNDTASYKCETQNPVSARRSDSVILNVLYGPDA
                     SGENLNLSCHAASNPPAQYSWFVNGTFQQSTQELFIPNITVNNSGSY
                     NRTTVTTITVYAEPPKPFITSNNSNPVEDEDAVALTCEPEIQNTTYL
                     PRLQLSNDNRTLTLLSVTRNDVGPYECGIQNELSVDHSDPVILNVLY
                     TYYRPGVNLSLSCHAASNPPAQYSWLIDGNIQQHTQELFISNITEKN
                     ASGHSRTTVKTITVSAELPKPSISSNNSKPVEDKDAVAFTCEPEAQN
                     LPVSPRLQLSNGNRTLTLFNVTRNDARAYVCGIQNSVSANRSDPVTL
                     SPPDSSYLSGANLNLSCHSASNPSPQYSWRINGIPQQHTQVLFIAKI
                     VSNLATGRNNSIVKSITVSASGTSPGLSAGATVGIMIGVLVGVALI"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2019;
  Best Local Similarity   72.7%;  Pred. No. 2.7e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              |||||||||||||||||||||     |  |    ||  |  | ||||
Db       1385 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1445 CCGTGGAGGACAAGGAT 1461


RESULT 15
A39900
LOCUS       A39900                2031 bp    DNA     linear   PAT
DEFINITION  Sequence 2 from Patent EP0618292.
ACCESSION   A39900
VERSION     A39900.1  GI:2296118
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
REFERENCE   1  (bases 1 to 2031)
  AUTHORS   Terskikh,A., Pelegrin,A. and Mach,J.
  TITLE     CEA derivatives lacking the Hydrophobic C-terminal domai
  JOURNAL   Patent: EP 0618292-A 2 05-OCT-1994;
            HOFFMANN LA ROCHE (CH)
COMMENT     Other publication CA 2116640 940926
            Other publication JP 6321992 941122.
FEATURES             Location/Qualifiers
     source          1. .2031
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2031;
```

```
ilarity   72.7%;  Pred. No. 2.7e-05;
 Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    | |    || |  | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1534

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1551



48                    2031 bp    DNA     linear   PAT 30-DEC-1997
ence 2 from patent US 5672513.
48
48.1  GI:2731283

own.
own.
assified.
bases 1 to 2031)
kikh,A., Pelegrin,A. and Mach,J.-P.
ino-embryonic antigen derivatives lacking the carboxyl terminal

nt: US 5672513-A 2 30-SEP-1997;
      Location/Qualifiers
      1. .2031
      /organism="unknown"
      /mol_type="unassigned DNA"


          56.4%;  Score 43.4;  DB 6;  Length 2031;
ilarity   72.7%;  Pred. No. 2.7e-05;
 Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    | |    || |  | ||||||||||
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1534

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1551



AASP                  2092 bp    mRNA    linear   PRI 19-JUL-1995
n mRNA for carcinoembryonic antigen pCEA80-11.
55
55.1  GI:29854
rnative splicing; carcinoembryonic antigen; cea gene; cell
sion molecule; cell surface protein.
 sapiens (human)
 sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 2092)
awa,S.
ect Submission
mitted (02-SEP-1989) Oikawa S., Suntory Institute for Biomedical
earch, 1-1-1 Wakayamadai, Shimamoto-cho, Mishima-gun Osaka 618,
an
hor describes alternative splicing which leads to a three base
r deletion (gca) in this CEA mRNA.
 also <M15042> for sequence of CEA mRNA.
      Location/Qualifiers
      1. .2092
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /tissue_type="colonic mucosa"
      CDS             <1. .>2092
                      /codon_start=2
                      /product="pCEA80-11 protein (647 AA)"
                      /protein_id="CAA34474.1"
                      /db_xref="GI:825638"
                      /db_xref="GOA:P06731"
                      /db_xref="SWISS-PROT:P06731"
                      /translation="SAPPHRWCIPWQRLLLTASLLTFWNPPTTAKLT
                      KEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATPGPAY
                      LLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNN
                      FTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDTA
                      ARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQY
                      TQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYEPPKPFIT
                      AVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRN
                      ELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPP
                      QQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPK
                      EDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFN
                      GIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLNLSCHSA
                      NGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSA
                      TVGIMIGVLVGVALI"
      misc_feature    947. .948
                      /note="site of 3 bp (gca) deletion"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 9;  Length 2092;
  Best Local Similarity   72.7%;  Pred. No. 2.7e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              |||||||||||||||||||     |    | |    || |  | ||||
Db       1461 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||
Db       1521 CCGTGGAGGACAAGGAT 1537


RESULT 18
A43169
LOCUS       A43169                 2097 bp    DNA     linear   PAT
DEFINITION  Sequence 35 from Patent WO9506067.
ACCESSION   A43169
VERSION     A43169.1  GI:2298545
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hor
REFERENCE   1  (bases 1 to 2097)
  AUTHORS   Bodmer,W.F., Durbin,H., Snary,D., Stewart,L.M., Young,S
            Bates,P.A.
  TITLE     MONOCLONAL ANTIBODIES FOR USE IN DIAGNOSIS AND TREATMEN
            COLORECTAL CANCER
  JOURNAL   Patent: WO 9506067-A 35 02-MAR-1995;
            IMP CANCER RES TECH (GB)
COMMENT     Other publication CA 2168440 950302
            Other publication AU 7390694 950321.
FEATURES             Location/Qualifiers
     source          1. .2097
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .2094
                     /note="unnamed protein product; Protein sequen
                     conflict with the conceptual translation"
                     /codon_start=1
                     /protein_id="CAA02706.1"
                     /db_xref="GI:2298546"
                     /db_xref="REMTREMBL:CAA02706"
                     /translation="SAPPHRWCIPWQRLLLTASLLTFWNPPTTAKL
                     KEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATPGPA
                     LLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSN
                     FTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDT
```

ARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGTFQQS
TQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEPPKPFITSNNSNPVEDE
DAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYECGIQ
NELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPPAQYSWLIDGN
IQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPKPSISSNNSKP
VEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFNVTRNDARAYV
CGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLNLSCHSASNPSPQYSWR
INGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSAG
ATVGIMIGVLVGVALI"

56.4%; Score 43.4; DB 6; Length 2097;
ilarity 72.7%; Pred. No. 2.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

:ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
:ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1522

:GTGGAGGACAAGGAT 77
:GTGGAGGACAAGGAT 1539

'9553 2097 bp DNA linear PAT 31-AUG-2000
ence 35 from patent US 5965710.
9553
9553.1 GI:10006297

own.
own.
assified.
bases 1 to 2097)
er,W.F., Durbin,H., Snary,D., Stewart,L.M.D., Young,S. and
s,P.A.
clonal antibodies for use in diagnosis and treatment of
rectal cancer
nt: US 5965710-A 35 12-OCT-1999;
Location/Qualifiers
1. .2097
/organism="unknown"
/mol_type="unassigned DNA"

56.4%; Score 43.4; DB 6; Length 2097;
ilarity 72.7%; Pred. No. 2.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1522

GTGGAGGACAAGGAT 77
GTGGAGGACAAGGAT 1539

3657 2106 bp DNA linear PAT 15-MAY-2001
ence 111 from Patent WO0130382.
3657
3657.1 GI:14139699

hetic construct
hetic construct
ficial sequences.

nstein,N., Tartaglia,J., Moingeon,P. and Barber,B.
od of inducing and/or enhancing an immune response to tumor
gens

JOURNAL Patent: WO 0130382-A 111 03-MAY-2001;
Aventis Pasteur Limited (CA)
FEATURES Location/Qualifiers
source 1. .2106
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
CDS 1. .2106
/note="unnamed protein product; modified CEA"
/codon_start=1
/transl_table=11
/protein_id="CAC38955.1"
/db_xref="GI:14139700"
/db_xref="REMTREMBL:CAC38955"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQAT
PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPS
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVT
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASN
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYEPP
EDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLL
GIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHA
DGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVS
SKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRT
AYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGADLNL
SWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKS
SAGATVGIMIGVLVGVALI"
ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2106;
Best Local Similarity 72.7%; Pred. No. 2.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
Db 1472 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1532 CCGTGGAGGACAAGGAT 1548

RESULT 21
AX192349
LOCUS AX192349 2106 bp DNA linear PAT
DEFINITION Sequence 3 from Patent WO0149317.
ACCESSION AX192349
VERSION AX192349.1 GI:15210326
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Emtage,P., Barber,B.H., Sambhara,S. and Sia,C.D.
TITLE Enhancing the immune response to an antigen by presensit
an inducing agent prior to immunizing with the inducing
the antigen
JOURNAL Patent: WO 0149317-A 3 12-JUL-2001;
Aventis Pasteur Limited (CA)
FEATURES Location/Qualifiers
source 1. .2106
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
CDS 1. .2106
/note="unnamed protein product; modified CEA"
/codon_start=1
/transl_table=11
/protein_id="CAC51319.1"
/db_xref="GI:15210327"
/db_xref="REMTREMBL:CAC51319"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPTT
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATP

PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNNSKPVEDK
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDTASYKCETQ
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGT
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYEPPKPFITSNNSNPV
EDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYEC
GIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPPAQYSWLI
DGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPKPSISSNN
SKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFNVTRNDAR
AYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGADLNLSCHSASNPSPQY
SWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGL
SAGATVGIMIGVLVGVALI"

56.4%; Score 43.4; DB 6; Length 2106;
milarity 72.7%; Pred. No. 2.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1531

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1548

93888 2106 bp DNA linear PAT 23-MAR-2002
uence 2 from Patent WO0210379.
93888
93888.1 GI:19701852

thetic construct
thetic construct
ificial sequences.

instein,N., Tartaglia,J., Tine,J.A., Panicali,D.L., Gritz,L. and
lom,J.
ified cea and uses thereof
ent: WO 0210379-A 2 07-FEB-2002;
ntis Pasteur Limited (CA) ; Therion Biologics (US) ; National
cer Institute (US)
Location/Qualifiers
1. .2106
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Cea modified polypeptide"

56.4%; Score 43.4; DB 6; Length 2106;
milarity 72.7%; Pred. No. 2.7e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1531

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1548

33977 2109 bp DNA linear PAT 15-MAY-2001
uence 4 from Patent WO0124832.
33977
33977.1 GI:14139917

thetic construct
thetic construct
artificial sequences.
REFERENCE 1
AUTHORS Pecher,G.
TITLE Pharmaceutical composition for treating and preventing
tumors, which express the tumor antigen mucin and/or th
carcinoembryonic antigen (ce a), and the use thereof
JOURNAL Patent: WO 0124832-A 4 12-APR-2001;
Pecher, Gabriele (DE)
FEATURES Location/Qualifiers
source 1. .2109
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Gen des humanen Carcinoembryonalen Anti
CDS 1. .2109
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC38973.1"
/db_xref="GI:14139918"
/db_xref="REMTREMBL:CAC38973"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQAT
PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPS
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVT
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASN
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEP
VEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTL
CGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCH
IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITV
NSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNR
RAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLN
YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVK
LSAGATVGIMIGVLVGVALI"
ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2109;
Best Local Similarity 72.7%; Pred. No. 2.7e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
|||||||||||||||||||||| | | || | | |||
Db 1475 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
||||||||||||||||
Db 1535 CCGTGGAGGACAAGGAT 1551

RESULT 24
AX468838
LOCUS AX468838 2109 bp DNA linear PAT
DEFINITION Sequence 12 from Patent WO0222803.
ACCESSION AX468838
VERSION AX468838.1 GI:21901578
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1
AUTHORS Chicz,R. and Tomlinson,A.
TITLE Peptide epitopes recognized by antigen specific cd4 + t
JOURNAL Patent: WO 0222803-A 12 21-MAR-2002;
Zycos Inc. (US)
FEATURES Location/Qualifiers
source 1. .2109
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2109;

```
nilarity   72.7%;  Pred. No. 2.7e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||      |     | |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1534

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1551



44683                2220 bp    DNA     linear   PAT 29-SEP-1999
ience 16 from patent US 5817512.
44683
44683.1  GI:5966148

iown.
iown.
lassified.
(bases 1 to 2220)
row,C.D., Porter,D.C. and Ansardi,D.C.
apsidated recombinant viral nucleic acid and methods of making
 using same
ent: US 5817512-A 16 06-OCT-1998;
      Location/Qualifiers
      1. .2220
      /organism="unknown"
      /mol_type="unassigned DNA"


          56.4%;  Score 43.4;  DB 6;  Length 2220;
ilarity   72.7%;  Pred. No. 2.7e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||      |     | |    ||  |  | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1630

GTGGAGGACAAGGAT 77
|||||||||||||||
GTGGAGGACAAGGAT 1647



2808                 2349 bp    DNA     linear   PAT 29-SEP-1999
ience 145 from patent US 5833975.
2808
2808.1  GI:5977670

iown.
iown.
assified.
bases 1 to 2349)
etti,E., Tartaglia,J. and Cox,W.I.
rypox virus expressing cytokine and/or tumor-associated antigen
 sequence
nt: US 5833975-A 145 10-NOV-1998;
      Location/Qualifiers
      1. .2349
      /organism="unknown"
      /mol_type="unassigned DNA"


          56.4%;  Score 43.4;  DB 6;  Length 2349;
ilarity   72.7%;  Pred. No. 2.7e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||||      |     | |    ||  |  | ||||||||||
```

```
Db       1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              ||||||||||||||||
Db       1718 CCGTGGAGGACAAGGAT 1734


RESULT 27
AR288121
LOCUS       AR288121                2349 bp    mRNA    linear   PAT
DEFINITION  Sequence 145 from patent US 6537594.
ACCESSION   AR288121
VERSION     AR288121.1  GI:31675400
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2349)
  AUTHORS   Paoletti,E., Tartaglia,J. and Cox,W.I.
  TITLE     Vaccina virus comprising cytokine and/or tumor associat
            genes
  JOURNAL   Patent: US 6537594-A 145 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .2349
                     /organism="unknown"
                     /mol_type="mRNA"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2349;
  Best Local Similarity   72.7%;  Pred. No. 2.7e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||||      |     | |    ||  |  | |||
Db       1658 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              ||||||||||||||||
Db       1718 CCGTGGAGGACAAGGAT 1734


RESULT 28
AR052807
LOCUS       AR052807                2434 bp    DNA     linear   PAT
DEFINITION  Sequence 144 from patent US 5833975.
ACCESSION   AR052807
VERSION     AR052807.1  GI:5977669
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2434)
  AUTHORS   Paoletti,E., Tartaglia,J. and Cox,W.I.
  TITLE     Canarypox virus expressing cytokine and/or tumor-associa
            DNA sequence
  JOURNAL   Patent: US 5833975-A 144 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .2434
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2434;
  Best Local Similarity   72.7%;  Pred. No. 2.7e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
              |||| ||||||||||||||||||      |     | |    ||  | |||||
Db       1655 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy         61 CCGTGGAGGACAAGGAT 77
              |||| ||||||||||||
Db       1715 CCGTGGAGGACAAGGAT 1731
```

```
:88120                 2434 bp    mRNA    linear   PAT 12-JUN-2003
quence 144 from patent US 6537594.
:88120
:88120.1  GI:31675399

nown.
nown.
lassified.
(bases 1 to 2434)
letti,E., Tartaglia,J. and Cox,W.I.
cina virus comprising cytokine and/or tumor associated antigen
es
ent: US 6537594-A 144 25-MAR-2003;
      Location/Qualifiers
      1. .2434
      /organism="unknown"
      /mol_type="mRNA"


          56.4%;  Score 43.4;  DB 6;  Length 2434;
milarity   72.7%;  Pred. No. 2.7e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    | |    || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1714

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1731



34671                 2888 bp    mRNA    linear   PRI 07-OCT-2003
o sapiens carcinoembryonic antigen-related cell adhesion
ecule 5, mRNA (cDNA clone MGC:34212 IMAGE:5184800), complete
:.
34671
34671.1  GI:21961633
.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 2888)
ausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
usner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
schul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
kins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
tchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
pleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
eetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
ninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
amson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
ernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
ley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
lalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
ey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
chez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
ffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
kson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
terfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
nerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
eration and initial analysis of more than 15,000 full-length
an and mouse cDNA sequences
c. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
88257
77932

REFERENCE   2  (bases 1 to 2888)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2002) National Institutes of Health,
            Gene Collection (MGC), Cancer Genomics Office, National
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN
            DNA Sequencing by: Baylor College of Medicine Human Ger
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lou
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Na
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information
            through the I.M.A.G.E. Consortium/LLNL at: http://image
            Series: IRAK Plate: 50 Row: e Column: 17
            This clone was selected for full length sequencing beca
            passed the following selection criteria: matched mRNA s
FEATURES             Location/Qualifiers
     source          1. .2888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:34212 IMAGE:5184800"
                     /tissue_type="Colon, Kidney, Stomach, adult, w
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2888
                     /gene="CEACAM5"
                     /note="synonyms: CEA, CD66e"
                     /db_xref="LocusID:1048"
                     /db_xref="MIM:114890"
     CDS             96. .2204
                     /codon_start=1
                     /product="CEACAM5 protein"
                     /protein_id="AAH34671.1"
                     /db_xref="GI:21961634"
                     /db_xref="LocusID:1048"
                     /translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT
                     VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQAT
                     PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPS
                     DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVT
                     NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASN
                     FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEP
                     VEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTL
                     CGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCH
                     IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITV
                     NSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNR
                     RAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLN
                     YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVK
                     LSAGATVGIMIGVLVGVALI"
     misc_feature    564. .755
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
     misc_feature    825. .1046
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
     misc_feature    1110. .1331
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
     misc_feature    1359. .1586
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
```

```
re     1632. .1865
       /note="IG; Region: Immunoglobulin"
       /db_xref="CDD:smart00409"
re     1917. .2072
       /note="IG; Region: Immunoglobulin"
       /db_xref="CDD:smart00409"


          56.4%;  Score 43.4;  DB 9;  Length 2888;
nilarity   72.7%;  Pred. No. 2.8e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
|||||||||||||||||||||      |    | |    ||  | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1629

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1646



530                  2928 bp    RNA     linear   PAT 29-SEP-1997
A encoding human carcinoembryonic antigen.
530
530.1  GI:2169883
1988177794-A/1.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
nalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 2928)
awa,S. and Nakazato,H.
AN CARCINOEMBRYONIC ANTIGEN
ent: JP 1988177794-A 1 21-JUL-1988;
TORY LTD
  Human
  JP 1988177794-A/1
  21-JUL-1988
  14-JAN-1987 JP 1987006851
  OIKAWA SHINZO, NAKAZATO HIROSHI

N15/00,C07H21/04,C07K13/00,C12P21/02,G01N33/574//G01N33/577, PC
12P21/02,
  C12R1:19);
  strandedness: Double;
  topology: Linear;
  hypothetical: No;
  anti-sense: No;
  *source: tissue_type=Colon tumor;
  *source: clone=pCEA 80-11;
  Key              Location/Qualifiers

  mRNA             1. .2929
  CDS              1. .2097
                   /product='human carcinoembryonic antigen' FT
_peptide     1. .60
                   /product='signal peptide of human  FT
     cacinoembryonic antigen'
  mat_peptide      61. .2094
                   /product='mature peptide of human  FT
     carcinoembryonic antigen'
  3'UTR            2098. .2929.
     Location/Qualifiers
     1. .2928
     /organism="Homo sapiens"
     /mol_type="genomic RNA"
     /db_xref="taxon:9606"


          56.4%;  Score 43.4;  DB 6;  Length 2928;
nilarity   72.7%;  Pred. No. 2.8e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |    |  |    ||  |  | |||
Db       1463 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1523 CCGTGGAGGACAAGGAT 1539


RESULT 32
E13123
LOCUS       E13123                  2928 bp    DNA     linear   PAT
DEFINITION  Human gene for carcinoembryonic antigen (CEA),partial c
ACCESSION   E13123
VERSION     E13123.1  GI:3251935
KEYWORDS    JP 1997132599-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1  (bases 1 to 2928)
  AUTHORS   Oikawa,S. and Nakazato,H.
  TITLE     CARCINOEMBRYONIC ANTIGEN PROTEIN AND ITS PRODUCTION
  JOURNAL   Patent: JP 1997132599-A 1 20-MAY-1997;
            SUNTORY LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 1997132599-A/1
            PD   20-MAY-1997
            PF   14-JAN-1987 JP 1996221770
            PI   OIKAWA SHINZO, NAKAZATO HIROSHI
            PC   C07K14/82,C12N15/09,C12P21/02//G01N33/574,G01N33/5
            strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2928
            FT                   /organism='Homo sapiens'
            FT                   /tissue_type='colon cancer'
            FT   CDS             <1. .2097
            FT                   /product='carcinoembryonic antigen
            FT   sig_peptide     <1. .90
            FT   mat_peptide     91. .2094
            FT                   /product='carcinoembryonic antigen
FEATURES             Location/Qualifiers
     source          1. .2928
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2928;
  Best Local Similarity   72.7%;  Pred. No. 2.8e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||      |    |  |    ||  |  | |||
Db       1463 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1523 CCGTGGAGGACAAGGAT 1539


RESULT 33
HUMCEA
LOCUS       HUMCEA                  2929 bp    mRNA    linear   PRI
DEFINITION  Human carcinoembryonic antigen mRNA.
ACCESSION   M15042
VERSION     M15042.1  GI:180198
KEYWORDS    Alu repeat; antigen; carcinoembryonic antigen; repeat r
SOURCE      Homo sapiens (human)
```

```
io sapiens
:aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 2929)
:awa,S., Nakazato,H. and Kosaki,G.
.mary structure of human carcinoembryonic antigen (CEA) deduced
)m cDNA sequence
)chem. Biophys. Res. Commun. 142 (2), 511-518 (1987)
.28144
.4146
.ginal source text: Human malignant colon tissue, cDNA to mRNA,
)nes pCEA[55-2,80-11].
:an copy of sequence [1] kindly provided by S.Oikawa
)-MAR-1987).
      Location/Qualifiers
      1. .2929
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="19q13.2"
      1. .2929
      /gene="CEA"
      <1. .2876
      /gene="CEA"
      /product="CEA mRNA"
      <1. .2099
      /gene="CEA"
      /note="carcinoembryonic antigen precursor"
      /codon_start=2
      /protein_id="AAA51963.1"
      /db_xref="GI:180199"
      /db_xref="GDB:G00-119-054"
      /translation="SAPPHRWCIPWQRLLLTASLLTFWNPPTTAKLTIESTPFNVAEG
      KEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATPGPAYSGREIIYPNAS
      LLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNNSKPVEDKDAVA
      FTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDTASYKCETQNPVS
      ARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGTFQQS
      TQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEPPKPFITSNNSNPVEDE
      DAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYECGIQ
      NELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPPAQYSWLIDGN
      IQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPKPSISSNNSKP
      VEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFNVTRNDARAYV
      CGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLNLSCHSASNPSPQYSWR
      INGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSAG
      ATVGIMIGVLVGVALI"
e     <1. .91
      /gene="CEA"
      /note="carcinoembryonic antigen signal peptide"
e     92. .2095
      /gene="CEA"
      /product="carcinoembryonic antigen"
ion   2215. .2462
      /note="truncated Alu repeat"
 bp upstream of PvuII site.

          56.4%;  Score 43.4;  DB 9;  Length 2929;
milarity   72.7%;  Pred. No. 2.8e-05;
  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||||     |    | |   ||  | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1523

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1540



32574                2974 bp    DNA     linear   PAT 09-JAN-2002
uence 3083 from Patent WO0194629.
32574
```

```
VERSION     AX332574.1  GI:18123208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endres
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening usi
            gene sets
  JOURNAL   Patent: WO 0194629-A 3083 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2974
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2974;
  Best Local Similarity   72.7%;  Pred. No. 2.8e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||     |    | |   ||  | | |||
Db       1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1649 CCGTGGAGGACAAGGAT 1665


RESULT 35
AX409670
LOCUS       AX409670                2974 bp    DNA     linear   PAT
DEFINITION  Sequence 2317 from Patent WO0229103.
ACCESSION   AX409670
VERSION     AX409670.1  GI:21442375
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2317 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .2974
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. M29540"
ORIGIN

  Query Match           56.4%;  Score 43.4;  DB 6;  Length 2974;
  Best Local Similarity   72.7%;  Pred. No. 2.8e-05;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;

Qy          1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
              ||||||||||||||||||||     |    | |   ||  | | |||
Db       1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy         61 CCGTGGAGGACAAGGAT 77
              |||||||||||||||||
Db       1649 CCGTGGAGGACAAGGAT 1665


RESULT 36
AX658324
LOCUS       AX658324                2974 bp    DNA     linear   PAT
```

quence 240 from Patent WO03000928.
58324
58324.1 GI:29160809

lo sapiens (human)
lo sapiens
:aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

lsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
ersen,M.W. and Elsner,H.I.
hods for identification of cancer cell surface molecules and
.cer specific promoters, and therapeutic uses thereof
ent: WO 03000928-A 240 03-JAN-2003;
n Medical A/S (DK)
Location/Qualifiers
1. .2974
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

56.4%; Score 43.4; DB 6; Length 2974;
milarity 72.7%; Pred. No. 2.8e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

CGTGGAGGACAAGGAT 77
CGTGGAGGACAAGGAT 1665

77147 2974 bp DNA linear PAT 27-MAR-2003
ience 340 from Patent WO02103028.
77147
77147.1 GI:29334600

) sapiens (human)
) sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
alia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

inova,A.V., Yankovsky,N.K., Kozlov,A.P., Lobashev,A.V. and
ovskaya,L.L.
silico screening for phenotype-associated expressed sequences
nt: WO 02103028-A 340 27-DEC-2002;
edical Center (RU)
Location/Qualifiers
1. .2974
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

56.4%; Score 43.4; DB 6; Length 2974;
ilarity 72.7%; Pred. No. 2.8e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

ACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
ACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

GTGGAGGACAAGGAT 77
GTGGAGGACAAGGAT 1665

AX805532
LOCUS AX805532 2974 bp DNA linear PA
DEFINITION Sequence 1 from Patent WO03059379.
ACCESSION AX805532
VERSION AX805532.1 GI:38522494
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho
REFERENCE 1
AUTHORS Klysner,S. and Voldborg,B.
TITLE Immunogenic CEA
JOURNAL Patent: WO 03059379-A 1 24-JUL-2003;
Pharmexa A/S (DK)
FEATURES Location/Qualifiers
source 1. .2974
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS 115. .2223
/codon_start=1
/product="CEA"
/protein_id="CAE75559.1"
/db_xref="GI:38522495"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQAT
PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPS
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVT
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASN
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEP
VEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTL
CGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCH
IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITV
NSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNR
RAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLN
YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVK
LSAGATVGIMIGVLVGVALI"
sig_peptide 115. .216
mat_peptide 217. .2157
/product="unnamed"
sig_peptide 2158. .2220
ORIGIN

Query Match 56.4%; Score 43.4; DB 6; Length 2974;
Best Local Similarity 72.7%; Pred. No. 2.8e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAAC
Db 1589 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAAC

Qy 61 CCGTGGAGGACAAGGAT 77
Db 1649 CCGTGGAGGACAAGGAT 1665

RESULT 39
HUMCEAF
LOCUS HUMCEAF 2974 bp mRNA linear PRI
DEFINITION Human carcinoembryonic antigen mRNA (CEA), complete cds.
ACCESSION M29540
VERSION M29540.1 GI:180222
KEYWORDS carcinoembryonic antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
REFERENCE 1 (bases 1 to 2974)
AUTHORS Barnett,T., Goebel,S.J., Nothdurft,M.A. and Elting,J.J.
TITLE Carcinoembryonic antigen family: characterization of cDN
for NCA and CEA and suggestion of nonrandom sequence var
their conserved loop-domains

omics 3 (1), 59-66 (1988)
22014
0478
ginal source text: Human carcinoma cDNA to mRNA, clone pc15LV7.
Location/Qualifiers
1. .2974
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="19q13.2"
1. .2974
/gene="CEA"
115. .2223
/gene="CEA"
/note="carcinoembryonic antigen"
/codon_start=1
/protein_id="AAA51967.1"
/db_xref="GI:180223"
/db_xref="GDB:G00-119-054"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPTTAKLTIESTPFN
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATPGPAYSGREIIY
PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNNSKPVEDK
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDTASYKCETQ
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGT
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEPPKPFITSNNSNP
VEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYE
CGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPPAQYSWL
IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITVSAELPKPSISSN
NSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFNVTRNDA
RAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLNLSCHSASNPSPQ
YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPG
LSAGATVGIMIGVLVGVALI"

56.4%; Score 43.4; DB 9; Length 2974;
nilarity 72.7%; Pred. No. 2.8e-05;
Conservative 0; Mismatches 21; Indels 0; Gaps 0;

GACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACTCCAAAC 60
||||||||||||||||||| | | | || | | ||||||||||
GACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACTCCAAAC 1648

CGTGGAGGACAAGGAT 77
|||||||||||||||
CGTGGAGGACAAGGAT 1665

ANTCE 3036 bp DNA linear PRI 08-AUG-1995
in carcinoembryonic antigen gene, complete cds.
103
103.1 GI:178676
cinoembryonic antigen.
o sapiens (human)
o sapiens
aryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
malia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(bases 1 to 3036)
achemin,N.
iblished (1987)
(bases 1 to 2541)
achemin,N., Benchimol,S., Cournoyer,D., Fuks,A. and
mers,C.P.
.ation and characterization of full-length functional cDNA
nes for human carcinoembryonic antigen
Cell. Biol. 7 (9), 3221-3230 (1987)
18876
)312
ginal source text: Human colonic adenocarcinoma cell line Ls180,
to mRNA, and DNA.
t entry and computer-readable sequence for [2],[1] kindly
vided by N.Beauchemin, 23-NOV-1987.

[1] revises [2].
FEATURES Location/Qualifiers
source 1. .3036
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="19q13.2"
mRNA 1. .2541
/product="CEA mRNA"
variation 58
/note="a in one clone; t in another"
/replace="t"
gene 97. .2205
/gene="CEA"
CDS 97. .2205
/gene="CEA"
/note="carcinoembryonic antigen precursor"
/codon_start=1
/protein_id="AAB59513.1"
/db_xref="GI:178677"
/db_xref="GDB:G00-119-054"
/translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPT
VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQAT
PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPS
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVT
NPVSARRSDSVILNVLYGPDAPTISPLNTSYRSGENLNLSCHAASN
FQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYAEP
VEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTL
CGIQNELSVDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCH
IDGNIQQHTQELFISNITEKNSGLYTCQANNSASGHSRTTVKTITV
NSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNR
RAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLN
YSWRINGIPQQHTQVLFIAKITPNNNGTYACFVSNLATGRNNSIVK
LSAGATVGIMIGVLVGVALI"
sig_peptide 97. .198
/gene="CEA"
/note="carcinoembryonic antigen signal peptide
mat_peptide 199. .2202
/gene="CEA"
/product="carcinoembryonic antigen"
repeat_region 2330. .2560
/note="Alu repeat"
ORIGIN 477 bp upstream of BglII site; chromosome 19.

Query Match 56.4%; Score 43.4; DB 9; Length 3036;
Best Local Similarity 72.7%; Pred. No. 2.8e-05;
Matches 56; Conservative 0; Mismatches 21; Indels 0;

Qy 1 AGACAATCACAGTCTCTGCGGAAGCATCATCCTCTGCATGGTCAGGTCATAACT
|||||||||||||||||||| | | | || | | ||||
Db 1571 AGACAATCACAGTCTCTGCGGAGCTGCCCAAGCCCTCCATCTCCAGCAACAACT

Qy 61 CCGTGGAGGACAAGGAT 77
|||||||||||||||||
Db 1631 CCGTGGAGGACAAGGAT 1647

Search completed: April 13, 2004, 19:12:09
Job time : 1110.5 secs